Heinz von Foerster

Sicht und Einsicht

Versuche zu einer operativen Erkenntnistheorie

Carl-Auer-Systeme Verlag

Carl-Auer Verlag

Sicht und Einsicht

Heinz von Foerster

Versuche zu einer operativen Erkenntnistheorie

2006

Über alle Rechte der deutschen Ausgabe verfügt Carl-Auer-Systeme
Verlag und Verlagsbuchhandlung GmbH Heidelberg
Fotomechanische Wiedergabe nur mit Genehmigung des Verlages
Umschlaggestaltung: WSP, Heidelberg
unter Verwendung eines Fotos von Thomas Reinagl
Foto: © Thomas Reinagl, Wien

Online Ausgabe 2006
ISBN 10: 3-89670-567-9
ISBN 13: 978-3-89670-567-0

© 1999, 2006 Carl-Auer-Systeme Verlag, Heidelberg
Autorisierte deutsche Fassung von Wolfram K. Köck

Bibliografische Information Der Deutschen Bibliothek
Die Deutsche Bibliothek verzeichnet diese Publikation
in der Deutschen Nationalbibliografie; detaillierte bibliografische
Daten sind im Internet über http://dnb.ddb.de abrufbar.

Informationen zu unserem gesamten Programm, unseren Autoren
und zum Verlag finden Sie unter: **www.carl-auer.de.**

Wenn Sie unseren Newsletter zu aktuellen Neuerscheinungen
und anderen Neuigkeiten abonnieren möchten, schicken Sie
einfach eine leere E-Mail an: **carl-auer-info-on@carl-auer.de.**

Carl-Auer Verlag
Häusserstr. 14
69115 Heidelberg
Tel. 0 62 21-64 38 0
Fax 0 62 21-64 38 22
E-Mail: info@carl-auer.de

Inhaltsverzeichnis

Quellenverzeichnis ... VIII

Vorbemerkungen zur deutschen Ausgabe ... IX

Zukunft der Wahrnehmung: Wahrnehmung der Zukunft ... 3
Epidemie ... 3
Prozeß-Substanz ... 4
Relation – Prädikat ... 5
Kastration ... 6
Die wissenschaftliche Methode ... 9
Qualität – Quantität ... 10
Dysgnosie ... 10
Trivialisierung ... 12
Zukunft ... 13

Die Verantwortung des Experten ... 17
1 Das „Ein-Hirn-Problem": die Wissenschaften vom Gehirn ... 21
2 Das „Zwei-Hirn-Problem": Erziehung ... 21
3 Das „Viel-Hirn-Problem": Gesellschaft ... 22
4 Das „All-Hirn-Problem": Menschheit ... 22

Über das Konstruieren von Wirklichkeiten ... 25
Das Postulat ... 25
Die Experimente ... 26
1 Der blinde Fleck ... 26
2 Skotom ... 26
3. Alternanten ... 27
4 Verstehen ... 27
Interpretation ... 29
Neurophysiologie ... 31
1 Evolution ... 31
2 Das Neuron ... 33
3 Die Übertragung ... 34
4 Die Synapse ... 34
5 Der Cortex ... 35
6 Descartes ... 36
7 Die Errechnung ... 36
8 Geschlossenheit ... 39
Bedeutung ... 40

Bibliothekare und Technik: eine Mesalliance? ... 43
Technik ... 43
Bibliothekare ... 44
Herausforderung ... 45
Konfusion ... 45
Kognition ... 46
Semantische Rechenprozesse ... 48
Computer für die Semantik ... 52
Ökonomische Aspekte ... 57

Kybernetik einer Erkenntnistheorie ... 65

Bemerkungen zu einer Epistemologie des Lebendigen ... 81
I Das Problem ... 81
II Einführung ... 82
III Bemerkungen ... 86

Gedanken und Bemerkungen über Kognition ... 95
Gedanken ... 95
Bemerkungen ... 100
Danksagung ... 112

Über selbst-organisierende Systeme und ihre Umwelten ... 115
Anhang ... 128

Gedächtnis ohne Aufzeichnung ... 133

Molekular-Ethologie: ein unbescheidener Versuch semantischer Klärung ... 173
I Einführung ... 173
II Theorie ... 176
A Allgemeine Vorbemerkungen ... 176
B Maschinen mit endlich vielen Zuständen ... 177
1 Deterministische Maschinen ... 177
2 Interagierende Maschinen ... 183
3 Probabilistische Maschinen ... 187
C Maschinen mit endlich vielen Funktionen ... 192
1 Deterministische Maschinen ... 192
2 Tesselierungen ... 194
III Biophysik ... 197
A Allgemeine Bemerkungen ... 197
B Molekularspeicher ... 200
C Molekularrechner ... 200
D Molekularträger ... 201
IV Zusammenfassung ... 203

Gegenstände: greifbare Symbole für (Eigen-)Verhalten 207

Anhang 214
Anhang A 214
Anhang B 216

Zitierte Literatur 217

Veröffentlichungen Heinz von Foersters 223

Namenverzeichnis 228

Sachwortverzeichnis 229

Quellenverzeichnis

(in der Reihenfolge des Inhalts)

1. Zukunft der Wahrnehmung: Wahrnehmungen der Zukunft
   "Perception of the Future and the Future of Perception", *Instructional Science* vol. 1 (1), 1972, 31–43.
2. Die Verantwortung des Experten
   "Responsibilities of Competence", *Journal of Cybernetics* vol. 2(2), 1972, 1–6.
3. Über das Konstruieren von Wirklichkeiten
   "On Constructing a Reality", in: *Environmental Design Research* ed. W. F. E. Preiser, vol. 2, Stroudsburg 1973, 35–46.
4. Bibliothekare und Technik: eine Mesalliance?
   "Technology – What Will it Mean to Librarians?", *Illinois Libraries* vol. 53(9), 1971, 785–803.
5. Kybernetik einer Erkenntnistheorie
   in: *Kybernetik und Bionik* hrsg. von W. D. Keidel/W. Händler/M. Spreng, München 1974, 27–46.
6. Bemerkungen zu einer Epistemologie des Lebendigen
   "Notes on an Epistemology for Living Things", *Biological Computer Laboratory Report No. 9.3*, University of Illinois, Urbana 1972.
7. Gedanken und Bemerkungen über Kognition
   "Thoughts and Notes on Cognition", in: *Cognition. A Multiple View* ed. P. L. Garvin, New York 1970, 25–48.
8. Über selbst-organisierende Systeme und ihre Umwelten
   "On self-organizing systems and their environments", in: *Self-Organizing Systems* eds. M. C. Yovits/S. Cameron, London 1960, 31–50.
9. Gedächtnis ohne Aufzeichnung
   "Memory without Record", in: *Learning, Remembering, and Forgetting*, vol. 1: *The Anatomy of Memory*, Palo Alto 1965, 388–433.
10. Molekular-Ethologie: ein unbescheidener Versuch semantischer Klärung
    "Molecular Ethology. An Immodest Proposal for Semantic Clarification", in: *Molecular Mechanisms in Memory and Learning* ed. G. Ungar, New York 1970, 213–248.
11. Gegenstände: Greifbare Symbole für (Eigen-)Verhalten
    "Objects: Tokens for Eigen-Behaviors", *Cybernetic Forum* vol. VIII, nos. 3 & 4, 91–96.

Einige einleitende Hinweise

Die in diesem Band versammelten Arbeiten Heinz von Foersters bedürfen keiner einführenden Kommentierung: sie sprechen klar und deutlich für sich wie für ihren Autor.

Das erstaunliche und außergewöhnliche Lebenswerk Heinz von Foersters entzieht sich jeder vorschnellen und schlagwortartigen Charakterisierung. Einem Außenstehenden – wie dem Hersteller dieser deutschen Versionen einiger seiner wichtigsten und interessantesten Arbeiten – steht dies schon gar nicht zu.

Sinnvoll und angebracht erscheint allerdings die Mitteilung einer autobiographischen Skizze, die Heinz von Foerster selbst veröffentlicht hat, sowie der Hinweis auf Würdigungen seines Lebenswerks von berufener Seite, von Weggefährten, Mitarbeitern und Sachkennern. Die vielfachen Pionierleistungen, die es erbracht hat, werden nämlich erst richtig deutlich, wenn seine historischen Verflechtungen – und sei es auch nur in Andeutungen – nachvollzogen werden.

Zunächst also in deutscher Fassung der größte – und nur geringfügig veränderte – Teil des autobiographischen Selbstporträts *en miniature*, das Heinz von Foerster 1964 veröffentlicht hat (*Current Anthropology* Bd. 5, S. 330). Es entstand, nachdem die *Wenner-Gren-Foundation for Anthropological Research* ihn zu ihrem Präsidenten gewählt hatte:

„Ich wurde – es war am 13. November 1911 – in eine lebensfrohe echte Wiener Familie germanisch-slawisch-jüdischer Abstammung hineingeboren. Zu meinen Vorfahren gehörten Künstler, Handwerker, Ingenieure, Bauern, Architekten und Juristen. Als Kind wollte ich unbedingt ‚Naturforscher' werden, und das bedeutete für mich eine romantische Mischung aus Fridtjof Nansen und Marie Curie. Ich träumte damals zwar davon, einmal sehr berühmt zu werden, war aber ein miserabler Schüler und machte nie meine Hausaufgaben. In den naturwissenschaftlichen Fächern hielt ich sie für lächerlich, und für die musischen hatte ich einfach keine Zeit. Es gab zuviel anderes zu tun – Schifahren, Bergsteigen, ‚Nachtarbeit' in einer Jazzcombo, und nicht zuletzt, zusammen mit einem meiner Cousins, das Basteln herrlicher neuer ‚Illusionen' für ein unübertreffliches Zauberkunststück. Der Beginn des Studiums setzte all dem ein Ende. Ich belegte Physik, denn wenn ich die Natur schon verstehen wollte, mußte ich wohl oder übel ihre Gesetze kennenlernen. An der Hochschule geriet ich nach kurzer Zeit in den Bann einer neuen Art von Magie, der des ‚Wiener Kreises', einer kleinen, aber energischen Gruppe von Wissenschaftsphilosophen. Wittgenstein, Schlick, Menger und Carnap beeindruckten mich gewaltig, und ich begann langsam, den Unterschied zwischen der Welt, wie sie ist, und ihrer symbolischen Repräsentation durch Sprache und Gleichungen zu begreifen. Ich mußte mehr wissen über die mysteriöse Beziehung zwischen dem Beobachter und dem Beobachteten: ich mußte den Beobachter verstehen!

Der Krieg zerstörte abrupt all diese hochfliegenden Pläne. Wieder aber rettete mich die Zauberkunst, und ich überlebte ohne Schaden an Leib und Seele. Während der Kriegsjahre arbeitete ich in verschiedenen Forschungslaboratorien für Plasmaphysik und Mikrowellenelektronik. Im Jahr 1945 kam ich nach Wien zurück, mit einer Frau, drei Kindern, geborgten Hemden, Hosen und Schuhen, um gleich am Aufbau des ersten Rundfunksenders der Nachkriegszeit mitzuarbeiten. Bis 1949 blieb ich als Programmdirektor für

Wissenschaft und Kunst dabei. In der gleichen Zeit half ich einer österreichischen Telefonfirma bei der Herstellung ihrer dringend benötigten technischen Ausstattung. Diese aufregenden Jahre voller Phasen der Erschöpfung, des Durcheinanders und doch wieder geistiger Verjüngung führten mich schließlich auch zurück zu meinem alten Rätsel: zum Problem des ‚Beobachters'. Unterstützt und ermutigt durch zwei großartige Männer, Viktor Frankl und Otto Pötzl, beide Fachärzte für Psychiatrie, veröffentlichte ich 1948 die Skizze einer quantenmechanischen Theorie des physiologischen Gedächtnisses.

Im Jahre 1949 besuchte ich Freunde in den Vereinigten Staaten: ich kam zum Abendessen, – und blieb für's Leben! Gleich nach meiner Ankunft hatte ich das Glück, Warren McCulloch kennenzulernen, der damals Professor für Neuropsychiatrie an der *University of Illinois* war, und der nicht nur die Daten für meine Theorie des Gedächtnisses hatte, sondern mich auch in diese Universität in Urbana einführte, wo mir erneut die Künste der Magie eine Stellung als Direktor des *Electron Tube Research Laboratory* im *Department of Electrical Engineering* eröffneten. Durch McCulloch und die von der *Josiah Macy jr. Foundation* finanzierten Fachtagungen über „Kybernetik: kreiskausal geschlossene und rückgekoppelte Mechanismen in biologischen und sozialen Systemen" lernte ich alle jene Leute kennen, die die theoretischen Grundlagen für die Erklärung des Verhaltens der ‚wirklich' komplexen, d.h. der teleologischen und selbstorganisierenden, Systeme entwickelten. Dazu gehörten Gregory Bateson, Julian Bigelow, Heinrich Klüver, Margaret Mead, Filmer Northrop, John von Neumann, Norbert Wiener, und viele andere. Damals konnte ich fast kein Englisch, und so übertrugen mir die Tagungsteilnehmer einfach die Redaktion der Kongreßberichte, damit ich die Sprache im Schnellverfahren lernen könnte.

Ich fand die Ideen, die aus diesen Tagungen erwuchsen, so faszinierend, daß ich nach sieben Jahren Forschung über Mikrowellenröhren und Ultrahochgeschwindigkeitsoszillographie um die Genehmigung eines Jahres Forschungsurlaub bat. Gewappnet mit einem sehr ehrenvollen Guggenheim-Fellowship konnte ich endlich in Ruhe darangehen, die Neurophysiologie jenes rätselhaften ‚Beobachters' genauer zu studieren. Nach einem Jahr Arbeit mit Warren McCulloch – inzwischen am *Massachusetts Institute of Technology* (M.I.T.) – und Arturo Rosenblueth in Mexico City kam ich nach Urbana zurück und konnte dank der Großzügigkeit meines *Department* 1957 das *Biological Computer Laboratory* (BCL) einrichten. Dort erforschten wir die Gesetze des Rechnens in lebenden Organismen. Probleme der Kognition, der Selbstreparatur und Selbstorganisation führten uns von Untersuchungen der Struktur geordneter Zellaggregate zur Analyse symbolischer Repräsentationen. Dabei habe ich gelernt, daß wir das einzelne Element nicht verstehen können, wenn wir es nicht als Teil eines großartigen Ganzen sehen. Wenn ich daher etwas über den ‚Beobachter', also den Menschen, wissen wollte, mußte ich ihn als Teil eines Ganzen sehen, und das bedeutet, als Teil der Menschheitsfamilie..."

Das legendäre *Biological Computer Laboratory* (BCL) der *University of Illinois* in Urbana blieb bis zum Jahre 1976 bestehen. Heinz von Foerster, Zauberkünstler, Forscher, Techniker und *homme du monde* in einem, war als sein Direktor zwei Jahrzehnte lang unermüdlicher *spiritus rector* eines – im Rückblick phantastisch anmutenden – internationalen Teams wissenschaftlicher und künstlerischer Avantgardisten, von denen hier nur W. Ross Ashby, Gotthard Günther, Lars Löfgren, Herbert Brün, Gordon Pask und Humberto R. Maturana genannt seien. (Ein Sammelband mit ausgewählten Arbeiten Maturanas ist bereits 1982 in dieser Reihe auf Deutsch erschienen: *Erkennen: Die Organisation und Verkörperung von Wirklichkeit.*) Die Arbeiten des BCL galten verschiedensten Problembereichen: Kybernetik, Kognition, Wahrnehmung, Lernen, Gedächtnis, Allgemeine Systemtheorie, Regelungstheorie und Regelungstechnik, biologische und soziale Systeme, transklassische mehrwertige Logik, Computerwissenschaft, semantische Netzwerke, rela-

tionale Datenstrukturen, Informationstechnik, Automatentheorie, Philosophie, Linguistik, Bewegungsnotation, Kybernetik 2. Ordnung (oder „Kybernetik der Kybernetik", d.h. Erforschung beobachtender Systeme), u.a.m. Die Ergebnisse – 14 000 Druckseiten – wurden nach der Emeritierung von Foersters, mit der das BCL geschlossen wurde, auf 146 Mikrofiches veröffentlicht (genaue Angaben s.u. S. 227). Bis heute sind weder der Ideenreichtum noch die konkreten Ergebnisse dieser Arbeiten hinreichend zur Kenntnis genommen, geschweige denn ausgeschöpft und fruchtbar gemacht worden!

Abschließend der Verweis auf drei ausführlichere Würdigungen des Lebens-Werks Heinz von Foersters:

1. *Cybernetics Forum*, das Publikationsorgan der *American Society for Cybernetics* (ASC), enthält im Band 9 (1979) eine Festschrift für Heinz von Foerster mit Beiträgen von Stuart Umpleby, Stafford Beer, Gordon Pask, Humberto R. Maturana, Lars Löfgren, Edwin Schlossberg und Kenneth L. Wilson. Dieses Sonderheft bietet neben ausführlicher dokumentarischer und bibliographischer Information eine Fülle persönlicher Erinnerungen an die gemeinsame Arbeit im BCL.

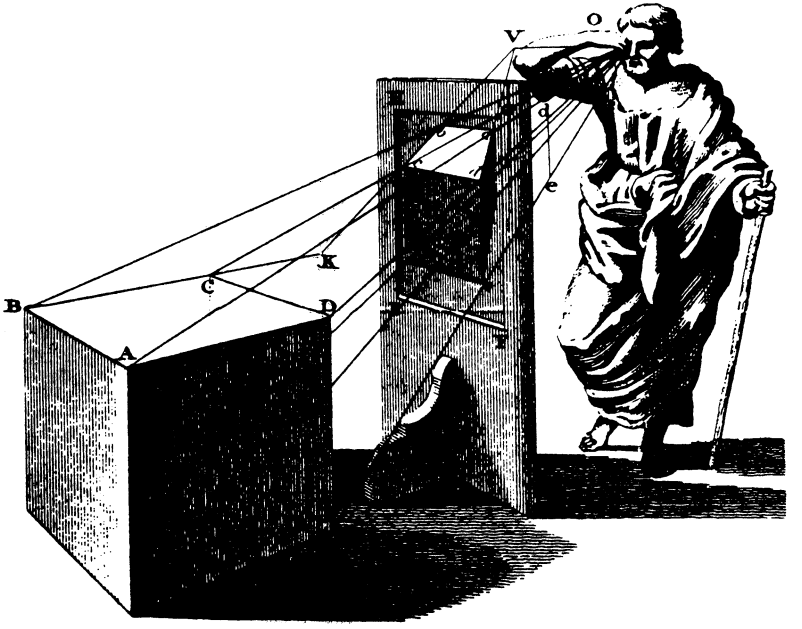
2. Im gleichen Jahr 1979 veröffentlichte auch Bernard Scott eine konzise Würdigung des wissenschaftlichen Lebenswerks von Heinz von Foerster im *International Cybernetics Newsletter* (ICNL), 12. Quartal, S. 209–214. Titel: „Heinz von Foerster – An Appreciation".

3. Eine knappe Würdigung („The Ages of Heinz von Foerster", S. xi–xvi) hat auch Francisco J. Varela seinem Sammelband mit 13 Arbeiten von Heinz von Foerster vorangestellt, der unter dem Titel *Observing Systems* 1982 in Seaside/California bei *Intersystems Publications* erschienen ist.

Zum Stand der Verbreitung und Diskussion der Arbeiten von und um Heinz von Foersters „Kybernetik und Kybernetik" vor allem in der Bundesrepublik Deutschland sei auf die „Vorbemerkungen" im oben angeführten deutschen Sammelband von H. R. Maturana (s. 11ff.) verwiesen.

Siegen, im Juni 1985 W. K. K.

Zukunft der Wahrnehmung: Wahrnehmung der Zukunft


V
O
B
C
K
D
A
d
e

Zukunft der Wahrnehmung: Wahrnehmung der Zukunft *

Kurzfassung

„Die Definition eines Problems sowie die zu dessen Lösung unternommenen Maßnahmen hängen weitgehend von der Sichtweise ab, in der die Individuen oder Gruppen, die das Problem entdeckt haben, jenes System auffassen, mit dem das Problem zusammenhängt. Ein Problem kann daher etwa definiert werden als ein mangelhaft interpretierter Output, oder als fehlerhafter Output einer fehlerhaften Outputvorrichtung, oder auch als fehlerhafter Output eines zwar fehlerfreien, aber gestörten Systems, oder schließlich als zwar richtiger, jedoch unerwünschter Output eines fehlerfreien und folglich unerwünschten Systems. Alle diese Definitionen außer der letzten verlangen korrigierendes Eingreifen, die letzte Definition allein aber fordert Veränderung und bietet somit ein unlösbares Problem für alle jene, die Veränderung ablehnen." (H. Brün 1971)

Gemeinplätze haben den fatalen Nachteil, daß sie durch Abstumpfen unserer Sinne die Wahrheit verschleiern. Kaum ein Mensch wird in Aufregung geraten, wenn er hört, daß in Zeiten der Kontinuität Zukunft und Vergangenheit gleich sind. Nur wenigen wird zum Bewußtsein kommen, was daraus folgt: in Zeiten soziokulturellen Wandels wird die Zukunft *nicht* sein wie die Vergangenheit.

Wenn wir uns aber gar kein klares Bild von der Zukunft machen, dann können wir auch nicht wissen, was wir tun sollen, da eines jedenfalls gewiß ist: Wenn wir selbst nicht handeln, wird mit uns gehandelt werden. Wenn wir also lieber Subjekte als Objekte sein wollen, dann muß unsere gegenwärtige Weltsicht, unsere Wahrnehmung also, auf die Zukunft gerichtet sein, nicht auf die Vergangenheit.

Epidemie

Meine Kollegen und ich erforschen gegenwärtig die Geheimnisse menschlichen Denkens und Wahrnehmens. Wenn wir von Zeit zu Zeit durch die Fenster unseres Labors auf die Ereignisse dieser Welt schauen, dann bedrückt uns immer mehr, was wir beobachten müssen. Die Welt scheint im Griff einer sich rasch ausbreitenden Krankheit zu sein, die schon fast globale Ausmaße erreicht hat. Im Individuum manifestieren sich die Symptome dieser Störung als ein fortschreitender Wahrnehmungsverfall, und eine verkommene Sprache ist der Infektionserreger, der diese Krankheit so hochgradig ansteckend macht.

Schlimmer noch: in fortgeschrittenen Stadien des Leidens werden die davon Befallenen völlig empfindungslos, verlieren Schritt für Schritt das Bewußtsein für ihr Gebrechen.

---

* Überarbeitete Fassung einer Rede zur Eröffnung der 24. "Annual Conference on World Affairs", University of Colorado, Boulder, Colorado/USA, vom 29. März 1971.

Diese Sachlage macht deutlich, warum mir unsere Wahrnehmungsfähigkeit Sorgen bereitet, wenn ich über die Zukunft nachdenke, denn:

Wenn wir nicht wahrnehmen können,
können wir die Zukunft nicht erkennen.
Wir wissen daher nicht, was jetzt zu tun ist.

Ich möchte behaupten, daß man dieser Schlußfolgerung durchaus zustimmen kann. Schaut man sich um, erscheint die Welt wie ein Ameisenhaufen, dessen Bewohner jeden Orientierungssinn verloren haben. Sie rennen ziellos herum, reißen einander in Stücke, beschmutzen ihr Nest, fallen über ihre Jungen her, investieren gewaltige Energien in den Bau komplizierter technischer Systeme, die nach Vollendung wieder aufgegeben werden oder dann, wenn sie weiter benutzt werden, die zuvor beobachtbare Zerrüttung nur noch vergrößern, usw. Die gezogenen Schlüsse scheinen somit den Tatsachen zu entsprechen. Sind aber die Prämissen gültig? Und was hat das alles mit Wahrnehmung zu tun?

Lassen Sie mich einige semantische Fallstricke beseitigen, bevor wir fortfahren, denn – wie ich bereits vorhin sagte – verkommene Sprache ist der Infektionserreger. Ein paar eindeutige Perversionen fallen Ihnen sicher sofort ein: z. B. „Friedenssicherung" statt „Kriegsvorbereitung", „Schutzmaßnahme" statt „Aggression", „Nahrungsentzug" statt „Vergiftung von Menschen, Tieren und Pflanzen". Glücklicherweise haben wir eine gewisse Immunität gegen solche Zumutungen entwickelt, denn zu lange schon sind wir von der Werbung mit syntaktischen Mißgeburten gefüttert worden, so z. B. mit „X ist besser", ohne daß je „als was" gesagt würde. Es gibt jedoch viele weit tiefer liegende semantische Konfusionen, und auf solche möchte ich jetzt Ihre Aufmerksamkeit lenken.

Es gibt drei Paare von Begriffen, bei denen ständig das eine Element eines Paars für das jeweils andere eingesetzt und so die Reichhaltigkeit unserer Vorstellungen beschnitten wird. Es ist zu einer Selbstverständlichkeit geworden, Prozeß und Substanz, Relation und Prädikat, und schließlich Qualität und Quantität zusammenzuwerfen. Ich möchte dies mit einigen wenigen Beispielen aus einem außerordentlich großen Katalog illustrieren und gleichzeitig zeigen, zu welch paralytischem Verhalten eine derartige begriffliche Dysfunktion führen kann.

Prozeß-Substanz

Die ursprünglichsten und zutiefst persönlichen Prozesse in jedem Menschen, und in der Tat in jedem Organismus, nämlich „Information" und „Erkenntnis", werden gegenwärtig durchwegs als Dinge bzw. Güter aufgefaßt, also als Substanzen. Information ist natürlich der Prozeß, durch den wir Erkenntnis gewinnen, und Erkenntnis sind die Prozesse, die vergangene und gegenwärtige Erfahrungen integrieren, um neue Tätigkeiten auszubilden, entweder als Nerventätigkeit, die wir innerlich als Denken und Wollen wahrnehmen können, oder aber als äußerlich wahrnehmbare Sprache und Bewegung (Maturana 1970a; 1970b; von Foerster 1969, 1970a).

Keiner dieser Prozesse kann „weitergegeben werden", wie man uns immer wieder sagt, z. B. mit Sätzen wie „... Universitäten sind Horte des Wissens, das von Generation zu Generation weitergegeben wird ..." usw., denn *Ihre* Nerventätigkeit ist ausschließlich *Ihre* Nerventätigkeit und – leider! – nicht *meine*.

Es ist kein Wunder, daß ein Bildungssystem, welches den Prozeß der Erzeugung neuer Prozesse mit der Verteilung von Gütern, genannt „Wissen", verwechselt, in den dafür bestimmten Empfängern große Enttäuschung hervorrufen muß, denn die Güter kommen nie an: es gibt sie nicht!

Die Konfusion, die Wissen als Substanz auffaßt, geht historisch auf ein Flugblatt zurück, das im 16. Jahrhundert in Nürnberg gedruckt wurde. Es zeigt einen sitzenden Schüler mit einem Loch im Kopf, in dem ein Trichter steckt. Daneben steht der Lehrer, der einen Kübel „Wissen" in den Trichter gießt: Buchstaben des Alphabets, Zahlen und einfache Gleichungen. Was die Erfindung des Rades für die ganze Menschheit gebracht hat, brachte der Nürnberger Trichter für die Bildung: es kann nun noch schneller abwärts gehen.

Gibt es ein Heilmittel? Natürlich, es gibt eines! Wir müssen Vorträge, Bücher, Diapositive, Filme usw. nicht als *Information*, sondern als *Träger* potentieller Information ansehen. Dann wird uns nämlich klar, daß das Halten von Vorträgen, das Schreiben von Büchern, die Vorführung von Diapositiven und Filmen usw. kein Problem löst, sondern ein Problem erzeugt: nämlich zu ermitteln, in welchen Zusammenhängen diese Dinge so wirken, daß sie in den Menschen, die sie wahrnehmen, neue Einsichten, Gedanken und Handlungen erzeugen.

Relation – Prädikat

Die Vermischung von Relationen und Prädikaten ist zu einem politischen Zeitvertreib geworden. In der Aussage „Spinat ist grün" ist „grün" ein Prädikat, in der Aussage „Spinat ist gut" bedeutet „gut" eine Relation zwischen der Chemie des Spinats und dem Beobachter, der den Spinat genießt. Er kann die Relation zwischen sich selbst und dem Spinat als „gut" bezeichnen. Unsere Mütter, die ersten Politiker, denen wir begegnen, machen sich die semantische Mehrdeutigkeit des syntaktischen Operators „ist" zunutze, indem sie uns sagen „Spinat *ist* gut", so als ob sie sagten „Spinat ist *grün*".

Wenn wir älter werden, werden wir mit solchen semantischen Verdrehungen überschüttet, die sicher lustig wären, wenn sie nicht so tiefgreifende Folgen hätten. Aristophanes hätte sehr gut eine Komödie schreiben können, in der die weisesten Menschen seines Landes sich eine Aufgabe stellen, die prinzipiell nicht bewältigt werden kann. Sie wollen ein für allemal die Eigenschaften feststellen, die einen obszönen Gegenstand oder eine obszöne Handlung definieren. Natürlich, „Obszönität" ist keine Eigenschaft, die den Dingen selbst angehört, sondern eine Subjekt-Objekt-Beziehung, denn wenn wir Herrn X ein Gemälde zeigen und er dieses obszön nennt, dann wissen wir eine Menge über Herrn X, aber sehr wenig über das Gemälde. Wenn also unsere Gesetzgeber endlich mit der von ihnen ausgedachten Liste von obszönen Eigenschaften an die Öffentlichkeit treten, werden wir eine Menge über die Gesetzgeber erfahren, ihre Gesetze aber werden gefährlicher Unsinn sein.

„Ordnung" ist ein weiterer Begriff, den wir, so trichtert man uns ein, in den Dingen selbst sehen sollen und nicht in unserer Wahrnehmung der Dinge. In den zwei Folgen A und B

A: 1, 2, 3, 4, 5, 6, 7, 8, 9
B: 8, 3, 1, 5, 9, 6, 7, 4, 2

erscheint die Folge A geordnet, die Folge B dagegen völlig durcheinander, bis man uns sagt, daß B die gleiche wunderschöne Ordnung aufweist wie A, denn B ist alphabetisch geordnet (acht, drei, eins ...). „Alles hat seine Ordnung, sobald man es versteht", sagt einer meiner Freunde, ein Neurophysiologe, der Ordnung dort sieht, wo ich nur einen völlig verworrenen Haufen von Zellen zu erkennen vermag. Wenn ich hier darauf bestehe, „Ordnung" als eine Subjekt-Objekt-Relation anzusehen und sie nicht mit einer Eigenschaft von Dingen zu verwechseln, dann mag das zu pedantisch erscheinen. Wenn es jedoch zum Problem von „Recht und Ordnung" kommt, kann eine derartige Verwechslung tödliche Konsequenzen haben. „Recht und Ordnung" ist kein Problem, es ist ein uns allen gemeinsames Ziel; das Problem ist nämlich, „*welches* Recht und *welche* Ordnung" gelten sollen, oder in anderen Worten, das Problem ist „Gerechtigkeit und Freiheit".

Kastration

Man kann solche Konfusionen in dem Glauben zur Seite wischen, daß sie ohne Schwierigkeiten zu korrigieren sind. Man kann behaupten, daß ich eben das gerade getan habe. Ich fürchte jedoch, daß dem nicht so ist. Die Wurzeln reichen tiefer als wir glauben. Wir scheinen in einer Welt aufgewachsen zu sein, die wir eher durch die Beschreibungen anderer sehen als durch unsere eigene Wahrnehmung. Dies hat zur Konsequenz, daß wir, statt die Sprache als ein Werkzeug zu benutzen, mit dem wir Gedanken und Erfahrungen ausdrücken, die Sprache als ein Werkzeug ansehen, das unsere Gedanken und unsere Erfahrungen festlegt.

Es ist nicht leicht, diese Behauptung zu beweisen, denn dafür ist nicht weniger erforderlich, als in unseren Kopf hineinzusteigen und die semantische Struktur bloßzulegen, die die Art und Weise unseres Wahrnehmens und Denkens widerspiegelt. Es gibt jedoch neue und faszinierende Experimente, aus denen diese semantischen Strukturen erschlossen werden können. Lassen Sie mich ein solches Experiment beschreiben, das meine Behauptung verdeutlicht.

Die von George Miller (1967) vorgeschlagene Methode besteht darin, mehrere Versuchspersonen unabhängig voneinander zu bitten, eine Reihe von Wörtern, die auf Karten aufgedruckt sind (Abbildung 1), nach ihrer Bedeutungsähn-

AGAIN	AIR	APPLE	BRING	CHEESE	COLD
COME	DARK	DOCTOR	EAT	FIND	FOOT
HARD	HOUSE	INVITE	JUMP	LIVE	MILK
NEEDLE	NOW	QUICKLY	SADLY	SAND	SEND
SLEEP	SLOWLY	SOFT	SUFFER	SUGAR	SWEET
TABLE	TAKE	VERY	WATER	WEEP	WHITE

Bild 1 Eine Auswahl von 36 auf Karten gedruckten Wörtern, die nach ihrer Bedeutungsähnlichkeit klassifiziert werden sollen.

lichkeit zu klassifizieren. Jede Versuchsperson kann so viele Klassen bilden, wie sie will, und jede beliebige Anzahl einzelner Wörter kann in eine Klasse aufgenommen werden. Die so gesammelten Daten können durch einen „Baum" so dargestellt werden, daß mit wachsender Entfernung der Verzweigungspunkte von der „Wurzel" des Baums die Übereinstimmung zwischen den Versuchspersonen zunimmt. So ergibt sich für die jeweilige Gruppe von Versuchspersonen ein Maß der Ähnlichkeit der Wortbedeutung.

Abbildung 2 zeigt die Resultate einer solchen "Cluster-Analyse" der 36 Wörter aus Abbildung 1 für 20 erwachsene Versuchspersonen („Wurzel" jeweils links).

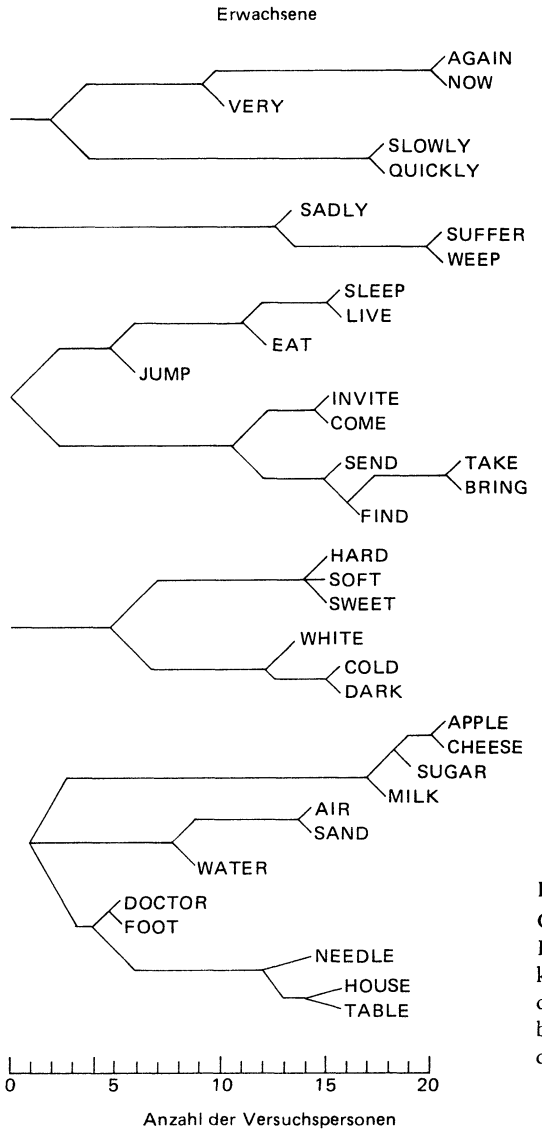


Bild 2

Clusteranalyse der 36 Wörter aus Bild 1, wie sie von 20 Erwachsenen klassifiziert wurden. Es fällt auf, daß syntaktische Kategorien genau beachtet, semantische Relationen dagegen fast völlig ignoriert wurden.

Ganz offensichtlich klassifizieren Erwachsene nach syntaktischen Kategorien, setzen Substantive in eine Klasse (unterer Baum), Adjektive in eine andere (nächsthöherer Baum), Verben in eine weitere, und schließlich ebenso jene kleinen Wörter, mit denen man nichts anzufangen weiß.

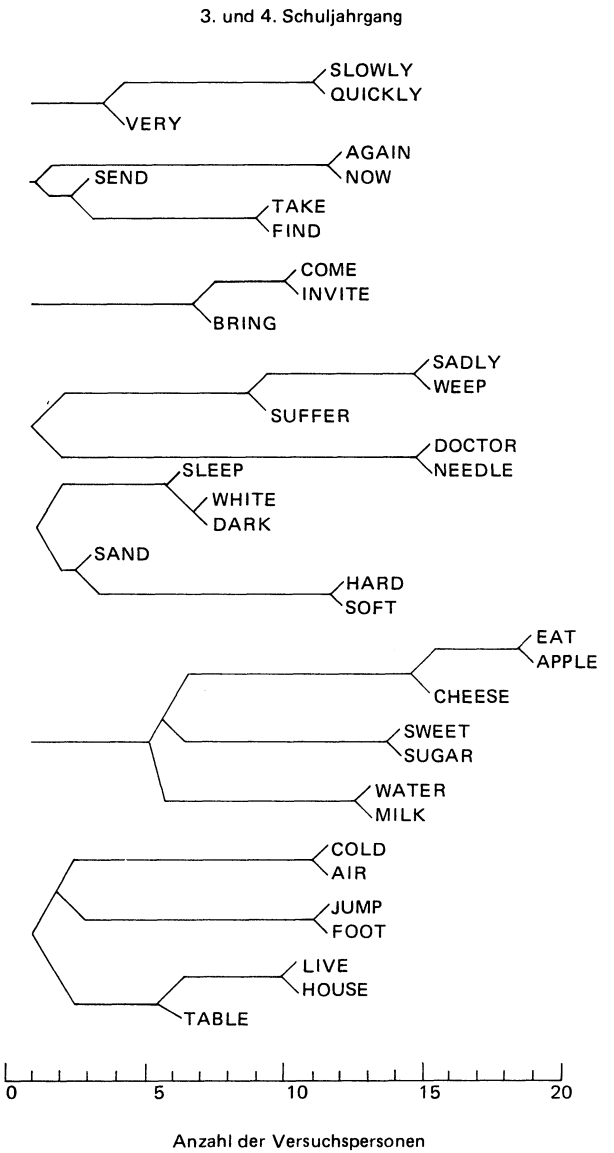


Bild 3 Die 36 Wörter der Bilder 1 und 2, wie sie von Kindern des dritten und vierten Schuljahres klassifiziert wurden. Es fällt auf, daß hier bedeutungsvolle kognitive Einheiten erscheinen, während syntaktische Kategorien fast völlig ignoriert werden.

Es zeigt sich ein eindrucksvoller Unterschied, wenn die Ergebnisse der Erwachsenen mit der Reichhaltigkeit der Wahrnehmung und bildhaften Phantasie von Kindern des 3. und 4. Schuljahrganges verglichen werden, die die gleiche Aufgabe gestellt bekamen (Abbildung 3). Miller kommentiert diese wunderschönen Ergebnisse so:

„Kinder neigen dazu, die Wörter zusammenzuordnen, mit denen man über die gleiche Sache spricht – was alle die sauberen syntaktischen Grenzen durchbricht, die für Erwachsene so wichtig sind. So stimmen alle 20 Kinder darin überein, das Verbum ‚essen' mit dem Substantiv ‚Apfel' zusammenzustellen; ‚Luft' ist für viele Kinder ‚kalt', sie brauchen den ‚Fuß', um zu ‚springen', sie ‚leben' in einem ‚Haus', und ‚Zucker' ist für sie ‚süß'; die Gruppe ‚Doktor', ‚Nadel', ‚leiden', ‚weinen', und ‚traurig' kann als besondere kleine Vignette angesehen werden."

Was stimmt mit unserer Erziehung nicht, daß sie die Beherrschung unserer Sprache dermaßen kastriert? Von den vielen dafür verantwortlichen Faktoren nenne ich nur den einen, der unsere Art zu denken in tiefgreifender Weise beeinflußt, nämlich die falsche Anwendung der „wissenschaftlichen Methode".

Die wissenschaftliche Methode

Die wissenschaftliche Methode ruht auf zwei tragenden Säulen:

1. Regeln, die in der Vergangenheit befolgt wurden, gelten auch für die Zukunft. Dies wird gewöhnlich als das Prinzip der Erhaltung von Regeln bezeichnet, und ich zweifle nicht daran, daß Ihnen allen dieses Prinzip vertraut ist. Die andere Säule jedoch steht im Schatten der ersten und ist daher nicht so klar erkennbar:
2. Fast alles im Universum soll irrelevant sein. Dies wird gewöhnlich das Prinzip der notwendigen und hinreichenden Ursache genannt. Was es fordert, wird unmittelbar einsichtig, wenn man erkennt, daß „Relevanz" eine triadische Relation ist, die eine Menge von Aussagen (P_1, P_2 ...) mit einer anderen Menge von Aussagen (Q_1, Q_2 ...) im Bewußtsein (B) eines Menschen verknüpft, der diese Verknüpfung herzustellen wünscht. Wenn P die Ursachen sind, die die wahrgenommenen Wirkungen Q erklären sollen, dann zwingt uns das Prinzip der notwendigen und hinreichenden Ursache dazu, unsere Wahrnehmung der Wirkungen immer weiter zu reduzieren, bis wir auf die notwendige und hinreichende Ursache stoßen, die die gewünschte Wirkung erzeugt: alles andere im Universum muß irrelevant bleiben.

Es ist leicht zu zeigen, daß die Fundierung der eigenen kognitiven Funktionen durch diese zwei Säulen für jede Betrachtung evolutionärer Prozesse kontraproduktiv ist, sei es das Aufwachsen eines Individuums oder die Veränderung einer ganzen Gesellschaft. Dies war schon Aristoteles bekannt, der zwei Arten von Ursachen unterschied, einmal die „Wirkursache", zum anderen die „Zweckursache", die uns zwei verschiedene Erklärungsmodelle für unbelebte Materie bzw. lebendige Organismen liefern, wobei der Unterschied darin liegt, daß die Wirkursache der Wirkung *vorausgeht*, während die Zweckursache der Wirkung *nachfolgt*. Wenn ich mit einem Streichholz die präparierte Fläche einer Streichholzschachtel reibe, dann ist dieses Reiben die (wirkende) Ursache dafür, daß sich das Streichholz entzündet. Die Ursache dafür jedoch, daß ich dieses Streichholz reibe, ist mein Wunsch, daß es brenne (Zweckursache).

Vielleicht erscheinen meine eingangs gemachten Bemerkungen durch diese Unterscheidung nun viel klarer. Natürlich dachte ich an die Zweckursache, als ich sagte, daß wir dann, wenn wir die Zukunft wahrnehmen können (das sich entzündende Streichholz), wissen, wie wir jetzt zu handeln haben (reiben!). Daraus will ich sofort eine Schlußfolgerung ziehen, nämlich:

In jedem Augenblick unseres Lebens sind wir frei, auf *die* Zukunft hin zu handeln, die wir uns wünschen.

Mit anderen Worten, die Zukunft wird so sein, wie wir sie sehen und erstreben. Dies kann nur für diejenigen ein Schock sein, die ihr Denken von dem Prinzip leiten lassen, daß für die Zukunft nur die Regeln gelten sollen, die in der Vergangenheit befolgt wurden. Für diese Menschen ist die Vorstellung einer „Veränderung" unbegreiflich, denn Veränderung ist der Prozeß, der die Regeln der Vergangenheit auslöscht.

Qualität – Quantität

Um die Gesellschaft vor den gefährlichen Folgen der Veränderung zu schützen, sind ganze Wirtschaftsunternehmen entstanden, und die Regierung hat mehrere Behörden eingerichtet, die sich damit beschäftigen, die Zukunft vorherzusagen, indem sie die Regeln der Vergangenheit anwenden. Es handelt sich um die Futurologen. Ihr Job besteht darin, Qualität und Quantität zu vermengen, und ihre Produkte sind sogenannte „Szenarios der Zukunft", in denen die Qualitäten gleichbleiben und nur die Quantitäten sich ändern: mehr Autos, breitere Straßen, schnellere Flugzeuge, größere Bomben usw. Obwohl diese „Szenarios der Zukunft" in einer sich verändernden Welt sinnlos sind, sind sie doch zu einem profitablen Geschäft für Unternehmen geworden, die sie an Firmen verkaufen, welche aus geplantem Produktverschleiß ihren Gewinn ziehen.

Mit der Diagnose der Unfähigkeit, qualitativen Wandel wahrzunehmen, d. h. einen Wandel unserer Subjekt-Objekt- und unserer Subjekt-Subjekt-Beziehungen, sind wir bis zur Wurzel der Epidemie vorgestoßen, von der ich in meinen einleitenden Bemerkungen gesprochen habe. Ein Beispiel aus der Neurophysiologie wird vielleicht die Unfähigkeit verständlich machen, die heutzutage auf der kognitiven Ebene wirksam ist.

Dysgnosie

Die visuellen Rezeptoren in der Retina, die Zapfen und die Stäbchen, arbeiten optimal nur unter bestimmten Belichtungsbedingungen. Über oder unter diesen Bedingungen vermindert sich die Sehschärfe oder die Farbdifferenzierung. Im Wirbeltierauge arbeitet die Retina jedoch fast immer unter optimalen Bedingungen, weil die Iris sich so verengt oder erweitert, daß bei wechselnden Helligkeitsbedingungen stets die gleiche Lichtmenge zu den Rezeptoren gelangt. Das vom Sehnerven „gesehene" Szenario weist daher stets die gleiche Helligkeit auf, unabhängig davon, ob wir uns in hellem Sonnenschein oder in einem abgedunkelten Zimmer befinden. Wie wissen wir also, ob es hell oder dunkel ist?

Die Information über diesen Tatbestand liegt in dem Regler, der die Aktivität des Sehnerven mit dem gewünschten Standard vergleicht und dafür sorgt, daß sich die Iris verengt, wenn diese Aktivität zu groß ist, und daß sie sich erweitert, wenn sie zu klein ist. Die Information über Helligkeit ergibt sich folglich nicht aus einer

Prüfung des Szenarios – dieses zeigt stets vergleichbare Helligkeit –, sie ergibt sich vielmehr aus einer Prüfung des Reglers, der die Wahrnehmung einer Veränderung unterdrückt.

Es gibt Menschen, die Schwierigkeiten haben, den Zustand ihres Reglers einzuschätzen, und die daher nur in geringem Maße fähig sind, verschiedene Helligkeitsgrade zu unterscheiden. Sie werden „Dysphotiker" genannt. Sie bilden den Gegensatz zu den Fotografen, die „Photiker" genannt werden könnten, denn sie haben einen besonders scharfen Sinn für Helligkeitsdiskriminierung. Es gibt nun aber Menschen, die Schwierigkeiten haben, jene Regler einzuschätzen, die ihre eigene Identität in einer sich wandelnden Welt erhalten. Ich möchte Individuen, die von dieser Störung betroffen sind, „Dysgnostiker" nennen, denn sie können sich selbst nicht erkennen. Da diese Störung schon ungewöhnlich weit verbreitet ist, hat man sie auch auf höchster nationaler Ebene bereits bemerkt.

Ich denke an jene hier in den Vereinigten Staaten, von denen man jetzt als der "Silent Majority" – der schweigenden Mehrheit – spricht. Ich aber sage, diese Menschen schweigen nicht, sie sind einfach stumm. Aber wie Sie wissen, sind es bei den Stummen nicht die Stimmbänder, sondern die Ohren, die nicht funktionieren: die Stummen sind taub! Das heißt, die schweigende Mehrheit kann nicht hören. Und sie kann nicht hören und kann nicht sehen, weil sie nicht hören und nicht sehen will, wie sich alles um sie verändert.

Der betrüblichste Aspekt dieser Beobachtung aber ist, daß auch ihr Gehör völlig in Ordnung ist. Sie könnten hören, wenn sie nur wollten: aber sie wollen nicht. Ihre Hörunfähigkeit ist ebenso freiwillig bzw. gewollt wie die Blindheit anderer.

Jetzt werden Sie natürlich Beweise für diese unerhörten Behauptungen verlangen. TIME-Magazine (1970) liefert sie in seinem Bericht über den Mittleren Westen der USA.

Da ist z. B. die Frau eines Rechtsanwalts in Glencoe/Illinois, die sich Sorgen macht über das Amerika, in dem ihre vier Kinder heranwachsen: „Ich möchte, daß meine Kinder in dem Amerika leben und aufwachsen, wie ich selbst es *erlebt* habe" – <Beachten Sie das Prinzip der Erhaltung von Regeln, demgemäß die Zukunft der Vergangenheit gleich ist!> –, in einem Amerika, in dem wir noch stolz waren, Bürger dieses Landes zu sein. Ich habe die Nase verdammt voll davon, mir all diesen Unsinn darüber anzuhören, wie schrecklich Amerika ist". <Beachten Sie die freiwillige bzw. gewollte Gehörlosigkeit!>

Ein anderes Beispiel ist der Zeitungsbibliothekar aus Pittsfield/Massachusetts, der sich über die Studentenunruhen ärgert: „Immer wenn ich Protestler sehe, sage ich, ‚Schau dir diese Krücken an!' ". <Beachten Sie die Verminderung der Sehschärfe!> „Aber dann sagt mir mein zwölfjähriger Sohn: ‚Das sind keine Krücken. Es ist ihr gutes Recht, das zu tun, was sie tun wollen.' " <Beachten Sie die un(v)-erwachsene Wahrnehmungsfähigkeit des Jungen!>

Die Tragödie, die in diesen Beispielen zum Ausdruck kommt, besteht darin, daß die Opfer der „Dysgnosie" nicht nur nicht wissen, daß sie nicht sehen, hören oder fühlen, sondern daß sie dies auch gar nicht wollen.

Wie können wir diese Situation ändern?

Trivialisierung

Ich habe bisher einige Beispiele für Wahrnehmungsstörungen gegeben, die unsere Sicht der Zukunft blockieren. Diese Symptome bilden insgesamt das Syndrom unserer epidemischen Erkrankung. Nun wäre der aber ein schlechter Arzt, der sich daran machte, den Patienten von einem Symptom nach dem anderen zu befreien, denn die Ausmerzung eines Symptoms kann ein anderes verschärfen. Gibt es einen gemeinsamen Nenner für die Wurzel des gesamten Syndroms?

Lassen Sie mich hier zwei Begriffe einführen, den Begriff der „trivialen" Maschine und den Begriff der „nicht-trivialen" Maschine. Der Ausdruck „Maschine" bezieht sich in diesem Zusammenhang auf wohldefinierte funktionale Eigenschaften einer abstrakten Größe, und nicht in erster Linie auf ein System von Zahnrädern, Knöpfen und Hebeln, obwohl solche Systeme jene abstrakten funktionalen Größen verwirklichen können.

Eine triviale Maschine ist durch eine eineindeutige Beziehung zwischen ihrem „Input" (Stimulus, Ursache) und ihrem „Output" (Reaktion, Wirkung) charakterisiert. Diese invariante Beziehung ist „die Maschine". Da diese Beziehung ein für allemal festgelegt ist, handelt es sich hier um ein deterministisches System; und da ein einmal beobachteter Output für einen bestimmten Input für den gleichen Input zu späterer Zeit ebenfalls gleich sein wird, handelt es sich dabei auch um ein vorhersagbares System.

Nicht-triviale Maschinen sind jedoch ganz andere Geschöpfe. Ihre Input-Output-Beziehung ist nicht invariant, sondern wird durch den zuvor erzeugten Output der Maschine festgelegt. Mit anderen Worten, ihre vorausgegangenen Arbeitsgänge legen ihre gegenwärtigen Reaktionen fest. Obwohl diese Maschinen auch deterministische Systeme sind, sind sie schon allein aus praktischen Gründen nicht vorhersagbar: ein einmal nach einem bestimmten Input beobachteter Output wird höchstwahrscheinlich zu späterer Zeit, auch wenn der Input gleich ist, ein anderer sein.

Um nun den grundlegenden Unterschied zwischen diesen beiden Arten von Maschinen zu begreifen, ist es hilfreich, sich „interne Zustände" in diesen Maschinen vorzustellen. Während in der trivialen Maschine in allen Fällen nur ein einziger interner Zustand an ihren internen Operationen mitwirkt, macht gerade der Übergang von einem internen Zustand zu einem anderen die nicht-triviale Maschine so unfaßbar.

Man kann nun diese Unterscheidung als die Analogie des 20. Jahrhundert zur Unterscheidung des Aristoteles zwischen den Erklärungsmodellen für unbelebte Materie und lebende Organismen ansehen.

Alle Maschinen, die wir konstruieren oder kaufen, sind hoffentlich triviale Maschinen. Ein Toaster sollte toasten, eine Waschmaschine waschen, ein Auto sollte in vorhersagbarer Weise auf die Handlungen seines Fahrers reagieren. Und in der Tat zielen alle unsere Bemühungen nur darauf, triviale Maschinen zu erzeugen, oder dann, wenn wir auf nicht-triviale Maschinen treffen, diese in triviale Maschinen zu verwandeln. Die Entdeckung der Landwirtschaft ist die Entdeckung, daß einige Aspekte der Natur trivialisiert werden können: Wenn ich heute pflüge, dann habe ich morgen Brot.

Zugegeben, in manchen Fällen gelingt uns die Herstellung idealer trivialer Maschinen nicht ganz. Eines Morgens etwa drehen wir den Zündschlüssel unseres

Autos, und das Miststück startet nicht. Offenbar hat es seinen internen Zustand in einer für uns undurchschaubaren Weise verändert, und zwar als Folge seiner vorhergegangenen Outputs (vielleicht hat es seinen Bezinvorrat aufgebraucht). Es hat so für einen Augenblick sein wahres Wesen als nicht-triviale Maschine enthüllt. Aber das ist natürlich eine unerhörte Sache und so ein Zustand muß sofort behoben werden.

Während nun unsere eifrigen Bemühungen um die Trivialisierung unserer Umwelt in einem Bereich nützlich und konstruktiv sein mögen, sind sie in einem anderen nutzlos und zerstörerisch. Trivialisierung ist ein höchst gefährliches Allheilmittel, wenn der Mensch es auf sich selbst anwendet.

Betrachten Sie etwa den Aufbau unseres Bildungssystems. Der Schüler kommt zur Schule als eine unvorhersagbare „nicht-triviale Maschine". Wir wissen nicht, welche Antwort er auf eine Frage geben wird. Will er jedoch in diesem System Erfolg haben, dann müssen die Antworten, die er auf unsere Fragen gibt, bekannt sein. Diese Antworten sind die „richtigen" Antworten:

F: „Wann wurde Napoleon geboren?"
A: „1769".
Richtig! (weil erwartet)
Schüler → Schüler
Aber:
F: „Wann wurde Napoleon geboren?"
A: „Sieben Jahre vor der amerikanischen Unabhängigkeitserklärung."
Falsch! (weil unerwartet)
Schüler → Nicht-Schüler.

Tests sind Instrumente, um ein Maß der Trivialisierung festzulegen. Ein hervorragendes Testergebnis verweist auf vollkommene Trivialisierung: der Schüler ist völlig vorhersagbar und darf daher in die Gesellschaft entlassen werden. Er wird weder irgendwelche Überraschungen noch auch irgendwelche Schwierigkeiten bereiten.

Zukunft

Ich nenne eine Frage, deren Antwort bekannt ist, eine „illegitime Frage". Wäre es nicht faszinierend, ein Bildungssystem aufzubauen, das von seinen Schülern verlangt, Antworten auf „legitime Fragen" zu geben, d. h. auf Fragen, deren Antworten unbekannt sind (H. Brün, persönliche Mitteilung)? Wäre es nicht noch faszinierender, sich eine Gesellschaft auszumalen, die ein solches Bildungssystem einrichten würde? Die notwendige Voraussetzung für diese Utopie ist, daß ihre Mitglieder einander als autonome und nicht-triviale Wesen auffassen. Eine derartige Gesellschaft wird, so sage ich vorher, einige ganz verblüffende Entdeckungen machen. Ich führe hier als Belege nur die folgenden drei auf:

1. „Bildung ist weder ein Recht, noch ein Privileg: sie ist eine Notwendigkeit."
2. „Bildung besteht darin, legitime Fragen stellen zu lernen."

Eine Gesellschaft, die diese beiden Entdeckungen gemacht hat, wird schließlich in der Lage sein, auch die dritte und utopischste zu machen:

3. „A geht es besser, wenn es B besser geht."

Wenn nun jemand in der heutigen Lage auch nur einen dieser drei Vorschläge ernsthaft machen wollte, wird er sicherlich in Schwierigkeiten geraten. Vielleicht erinnern Sie sich an die Geschichte, die Iwan Karamasow erfindet, um den Geist seines jüngeren Bruders Aljoscha ein bißchen zu kitzeln. Es ist die Geschichte vom Großinquisitor. Wie Sie wissen, geht der Großinquisitor an einem sehr angenehmen Nachmittag durch seine Stadt Sevilla. Er ist gut gestimmt. Am Morgen hat er auf dem Scheiterhaufen um die 120 Häretiker verbrannt, er hat seine Pflicht gut getan, alles ist in bester Ordnung. Plötzlich sieht er vor sich eine Menschenmenge, und er geht näher heran, um zu sehen, was da geschieht. Er sieht einen Fremden, der einem Lahmen seine Hand auflegt, und wie der Lahme auf einmal gehen kann. Dann bringt man ein blindes Mädchen zu ihm, der Fremde legt seine Hand auf ihre Augen, und auf einmal kann es sehen. Der Großinquisitor weiß sofort, wer Er ist, und er sagt zu seinen Henkern: „Nehmt diesen Menschen fest!" Sie stürzen sich auf ihn, packen ihn und werfen ihn ins Gefängnis. In der Nacht besucht der Großinquisitor den Fremden in seiner Zelle und sagt zu ihm: „Schau, ich weiß wer Du bist, Unruhestifter. Eintausendundfünfhundert Jahre hat es uns gekostet, den ganzen Unfrieden zu beseitigen, den Du gesät hast. Du weißt sehr wohl, daß Menschen selbständig keine Entscheidungen treffen können. Du weißt sehr wohl, daß Menschen nicht frei sein können. *Wir* müssen ihre Entscheidungen treffen. *Wir* sagen ihnen, wer sie sein sollen. Du weißt das sehr wohl. Daher werde ich Dich morgen auf dem Scheiterhaufen verbrennen." Der Fremde steht auf, umarmt den Großinquisitor und küßt ihn. Der Großinquisitor geht hinaus, aber als er die Zelle verläßt, schließt er die Tür nicht zu, und der Fremde verschwindet in die Dunkelheit der Nacht. ...

Lassen Sie uns an diese Geschichte denken, wenn wir solchen Unruhestiftern begegnen, und lassen Sie uns die Türen für sie offenhalten. Wir werden sie an einer schöpferischen Tat erkennen:

„Sehet!"

Und es ward Licht.

Die Verantwortung des Experten

Zorn
hochfart
neid
Tragkeit
fressery
vnkeuschait
Geitkat

Die Verantwortung des Experten *

Auf unserem letzten Jahrestreffen habe ich Ihnen ein Theorem vorgelegt, das Stafford Beer bei anderer Gelegenheit als „Heinz von Foersters Theorem Nr. 1" bezeichnet hat. Einige von Ihnen werden sich vielleicht noch daran erinnern, es lautet wie folgt:

„Je tiefer das Problem, das ignoriert wird, desto größer sind die Chancen, Ruhm und Erfolg einzuheimsen."

Kraft dieser auf einen einzigen Fall gegründeten Tradition möchte ich ihnen heute erneut ein Theorem vorlegen, das ich in aller Bescheidenheit als „Heinz von Foersters Theorem Nr. 2" bezeichnen will. Um es richtig formulieren zu können, muß ich eine historische Bemerkung voranschicken. Im Englischen ist es nicht wie im Deutschen seit Dilthey möglich, zwischen Natur- und Geisteswissenschaften zu unterscheiden. In England, wie in Amerika, spricht man von den „hard-" und den "soft sciences", eine Unterscheidung, die sicherlich von einem "hard scientist" erfunden wurde. Mit diesem Kommentar lautet mein Theorem Nr. 2 folgendermaßen:

„Die 'hard sciences' sind erfolgreich, weil sie sich mit den 'soft problems' beschäftigen; die 'soft sciences' haben zu kämpfen, denn sie haben es mit den 'hard problems' zu tun."

Sollten Sie bereit sein, sich die Sache genauer anzusehen, dann werden Sie sicherlich entdecken, daß Theorem Nr. 2 als Folgesatz zu Theorem Nr. 1 dienen könnte. Dies wird unmittelbar einsichtig, wenn wir für einen Augenblick die Untersuchungsmethode der "hard sciences" näher betrachten. Ist ein System zu komplex, um verstanden zu werden, dann wird es in kleinere Stücke zerlegt. Sind diese immer noch zu komplex, werden auch sie zerkleinert, und so geht es weiter, bis die Stücke schließlich so klein sind, daß zumindest eines davon verständlich ist. Das Wunderbare an diesem Prozeß, an der Methode der Reduktion, am „Reduktionismus", ist, daß sie unweigerlich zum Erfolg führen.

Leider befinden sich die "soft sciences" nicht in einer ähnlich glücklichen Lage. Denken wir etwa nur an die Soziologen, die Psychologen, Anthropologen, Linguisten usw. Würden sie die komplexen Systeme, mit denen sie sich befassen, also die Gesellschaft, die Psyche, die Kultur, die Sprache usw., in derselben Weise so reduzieren, daß sie sie zur immer genaueren Untersuchung in immer kleinere Teile zerlegen, dann könnten sie schon nach wenigen Schritten nicht mehr behaupten, daß sie es noch mit dem System zu tun haben, mit dem sie sich ursprünglich beschäftigen wollten. Dies liegt daran, daß diese Wissenschaftler es mit im wesentlichen nicht-linearen Systemen zu tun haben, deren kennzeichnende Eigenschaften in den *Interaktionen* zwischen dem bestehen, was man jeweils als die

* Überarbeitete Fassung des Grundsatzreferats zur Herbsttagung der American Society for Cybernetics am 9. Dezember 1971 in Washington/D.C.

„Teile" dieser Systeme auffaßt, während die Eigenschaften dieser „Teile" zum Verständnis des Funktionierens dieser Systeme als *Ganzes* wenig oder gar nichts beitragen. Wenn also ein Wissenschaftler, der in einer "soft science" arbeitet, im Gebiet seiner Wahl zu verbleiben wünscht, muß er mit einem riesigen Problem fertig werden: Er kann es sich nicht leisten, die tatsächliche Komplexität seines Systems aus den Augen zu verlieren, es wird aber von Tag zu Tag dringender, die sich ihm stellenden Probleme zu lösen. Und dies nicht bloß, weil er seinen Spaß daran hätte. Es ist inzwischen völlig klar geworden, daß seine Probleme uns alle angehen. „Entartung unserer Gesellschaft", „psychische Störungen", „kulturelle Erosion", „Versagen der Kommunikation" und all die vielen anderen „Krisen" unserer Zeit sind ebenso sehr unsere Probleme wie seine. Wie können wir zu ihrer Lösung beitragen?

Mein Vorschlag lautet, das *Fachwissen* – und nicht die Methode der Reduktion –, das wir in den Naturwissenschaften erworben haben, zur Lösung der harten Probleme in den Geisteswissenschaften einzusetzen. Ich füge sofort hinzu, daß dieser Vorschlag überhaupt nicht neu ist. Ich lege aber hiermit die These vor, daß es die *Kybernetik* ist, die das harte Fachwissen mit den harten Problemen der Geisteswissenschaften verknüpft. Diejenigen unter uns, die die frühe Entwicklung der Kybernetik miterlebt haben, werden sich sicherlich daran erinnern, daß unsere Wissenschaft als die Erforschung „kreis-kausal geschlossener und rückgekoppelter Mechanismen in biologischen und sozialen Systemen" verstanden wurde, bevor Norbert Wiener den Namen „Kybernetik" schuf, und daß sie noch Jahre nach der Abfassung seines berühmten Buches so beschrieben worden ist. Natürlich hat Norbert Wiener durch seine Definition der Kybernetik als der Wissenschaft von „Regelung und Signalübertragung im Lebewesen und in der Maschine" die Generalisierung dieser Begriffe weiter vorangetrieben, so daß „Kybernetik" heute die Wissenschaft der *Regelung* im allgemeinsten Sinne benennt.

Wenn sich nun unsere Wissenschaft mit diesem allgemeinen und allumfassenden Phänomen der Regelung befaßt, warum verfügt sie dann noch nicht wie die meisten unserer Schwesterdisziplinen über einen Schutzpatron oder eine Göttin, die uns auf der Suche nach neuen Erkenntnissen ihre Gunst schenken und unsere Gesellschaft gegen Übel von außen wie von innen schützen? Astronomen und Physiker werden von Urania betreut, Demeter schützt die Landwirtschaft, und die Musen helfen den verschiedenen Künsten und Wissenschaften. Wer aber hilft der Kybernetik?

Als ich eines nachts über diese kosmische Frage nachdachte, hatte ich ganz plötzlich eine Erscheinung. Leider war es keine der reizenden Göttinnen, die die anderen Künste und Wissenschaften beglücken. Das lustige kleine Geschöpf, das auf meinem Schreibtisch saß, mußte ein Dämon sein. Kurz darauf begann er auch zu reden. Ich hatte recht. „Ich bin Maxwells Dämon", sagte er. Und verschwand.

Als ich meine Fassung wiedergefunden hatte, war mir klar, daß nur dieser ehrenwerte Dämon unser Schutzpatron sein konnte, denn Maxwells Dämon ist *das Paradigma der Regelung*.

Wie Sie wissen, regelt Maxwells Dämon den Fluß der Moleküle zwischen zwei Behältern auf eine höchst *unnatürliche* Weise, nämlich so, daß Wärme vom kalten Behälter zum heißeren fließt. Im Gegensatz dazu fließt im natürlichen Ablauf der Ereignisse – also ohne den Eingriff des Dämons – Wärme immer vom heißen Behälter zum kälteren.

Sicherlich erinnern Sie sich auch, wie unser Dämon vorgeht. Er wacht über eine kleine Öffnung zwischen den beiden Behältern, die er freigibt, um ein schnelles Molekül aus dem kalten Bereich oder ein langsames aus dem heißen Bereich durchzulassen, die er aber ansonsten geschlossen hält. Durch dieses Manöver erreicht er, daß der kalte Behälter kälter und der heiße Behälter heißer wird, – was scheinbar den Zweiten Hauptsatz der Thermodynamik über den Haufen wirft. Natürlich haben wir inzwischen herausgefunden, daß der Zweite Hauptsatz unberührt bleibt, auch wenn der Dämon diesen perversen Wärmefluß tatsächlich bewerkstelligt. Er benötigt nämlich zur Feststellung der Geschwindigkeit der ankommenden Moleküle eine Taschenlampe. Wäre er im thermalen Gleichgewicht mit einem der Behälter, dann könnte er nämlich überhaupt nichts sehen: Er wäre Teil eines schwarzen Körpers. Da er seine Machinationen nur so lange betreiben kann, wie die Batterie seiner Taschenlampe vorhält, müssen wir in das System mit dem aktiven Dämon nicht nur die Energie der zwei Behälter, sondern auch die Energie der Batterie einbeziehen. Die durch das Nachlassen der Batterie gewonnene Energie wird nicht vollständig durch die Negentropie kompensiert, die durch die zunehmende Ungleichheit der beiden Behälter gewonnen wird.

Die Moral dieser Geschichte ist schlicht die, daß unser Dämon den Zweiten Hauptsatz der Thermodynamik nicht außer Kraft setzen, daß er aber durch seine Regelungsaktivität die Degradation der verfügbaren Energie, d.h. die Zunahme der Entropie, beliebig verlangsamen kann.

Dies ist nun in der Tat eine sehr wichtige Beobachtung, denn sie verdeutlicht die überragende Bedeutung von Regelungsmechanismen in lebenden Organismen. Lebewesen lassen sich daher als Manifestationen des Maxwellschen Dämons ansehen, die ständig die Degradation des Energieflusses, d.h. die Zunahme der Entropie verzögern. Mit anderen Worten, Organismen sind als Regelungssysteme „Entropieverzögerer".

Außerdem ist Maxwells Dämon, wie ich gleich zeigen werde, nicht nur ein Entropieverzögerer und ein Paradigma der Regelung, sondern auch funktional isomorph einer universalen Turingmaschine. Die drei Begriffe der Regelung, der Entropieverzögerung und des Rechnens bilden somit ein in sich verknüpftes Begriffsnetz, das für mich das Wesen der Kybernetik ausmacht.

Ich werde nun kurz meine Behauptung begründen, daß Maxwells Dämon nicht nur das Paradigma der Regelung, sondern auch des Rechnens ist.

Wenn ich den Begriff des „Rechnens" verwende, dann beschränke ich ihn nicht auf spezifische Operationen wie z.B. Addition, Multiplikation usw. Ich möchte „Rechnen" im allgemeinsten Sinn als einen Mechanismus oder „Algorithmus" des *Ordnens* bzw. der Erzeugung von *„Ordnung"* verstehen. Die ideale – oder vielleicht sollte man sagen: die allgemeinste – Darstellung eines solchen Mechanismus ist natürlich die Turingmaschine, und anhand dieser Maschine möchte ich einige meiner Behauptungen verdeutlichen.

Es gibt zwei Ebenen, auf denen wir von „Ordnen" sprechen können (L. Löfgren, 1967). Einmal wollen wir eine gegebene Anordnung von Gegenständen beschreiben, zum anderen gewisse Dinge entsprechend bestimmten Beschreibungen neu anordnen. Diese beiden Operationen sind in der Tat die Grundlage für all das, was wir „Rechnen" nennen.

Sei A eine bestimmte Anordnung. Diese Anordnung kann dann durch eine universale Turingmaschine aus einem geeigneten Anfangsausdruck auf ihrem Band berechnet werden, den wir als eine „Beschreibung" von A, D(A), bezeichnen wollen. Die Länge L(A) dieser Beschreibung hängt vom benutzten Alphabet (von der benutzten Sprache) ab. Wir können folglich feststellen, daß eine Sprache α_1 dann und nur dann in der Anordnung A mehr an Ordnung aufzeigt als eine andere Sprache α_2, wenn die Länge $L_1(A)$ der passenden Ausgangsbeschreibung auf dem Band zur Berechnung von A kleiner ist als $L_2(A)$ oder *mutatis mutandis.*

Dies gilt für die erste oben genannte Ebene und führt uns unmittelbar zur zweiten.

Unter all den geeigneten Anfangsbeschreibungen einer Anordnung A_1 auf dem Band gibt es eine kürzeste: $L^*(A_1)$. Wenn A_1 neu geordnet wird und somit A_2 ergibt, dann sei A_2 von höherer Ordnung als A_1 dann und nur dann, wenn die kürzeste Anfangsbeschreibung auf dem Band $L^*(A_2)$ kürzer ist als $L^*(A_1)$ oder *mutatis mutandis.*

Dies gilt für die zweite oben genannte Ebene und führt uns zu einer abschließenden Aussage über vollkommenes Ordnen (Rechnen).

Unter all den Anordnungen A_i gibt es eine Anordnung A^*, für die die geeignete Anfangsbeschreibung auf dem Band die kürzeste ist, nämlich $L^*(A^*)$.

Ich hoffe, daß diese Beispiele klargemacht haben, daß lebende Organismen – wir setzen sie nun an die Stelle der Turingmaschine –, die mit ihren Umwelten (Anordnungen) interagieren, über verschiedene Möglichkeiten verfügen:

1. Sie können „Sprachen" entwickeln (Sensoren, neuronale Kodes, motorische Organe usw.), die besser zu der jeweils gegebenen Umwelt „passen" (d.h. mehr an Ordnung erschließen).
2. Sie können ihre Umwelten so lange verändern, bis diese zu ihrer Konstitution „passen".
3. Sie können beides tun. Es sollte jedoch festgehalten werden, daß jede der von ihnen gewählten Möglichkeiten durch Rechnen verwirklicht wird. Ich habe also nun zu zeigen, daß solches Rechnen in der Tat der Aktivität unseres Dämons funktional isomorph ist.

Die entscheidende Funktion einer Turingmaschine läßt sich durch fünf Operationen bestimmen:

1. *Lies* das Inputsymbol x.
2. *Vergleiche* x mit z, dem inneren Zustand der Maschine.
3. *Schreibe* das passende Outputsymbol y.
4. *Verändere* den inneren Zustand z zum neuen Zustand z′.
5. *Wiederhole* die obige Folge mit einem neuen Inputzustand x′.

In ähnlicher Weise läßt sich die wesentliche Funktion des Maxwellschen Dämons durch fünf Operationen angeben, die den eben genannten äquivalent sind:

1. *Lies* die Geschwindigkeit v des ankommenden Moleküls M.
2. *Vergleiche* $(mv^2/2)$ mit der mittleren Energie $\langle mv^2/2\rangle$ (Temperatur T) etwa des kühleren Behälters (innerer Zustand T).
3. *Öffne* den Verschluß, wenn $(mv^2/2)$ größer ist als $\langle mv^2/2\rangle$; halte sonst die Öffnung geschlossen.
4. *Verändere* den inneren Zustand T zum neuen (kühleren) Zustand T′.
5. *Wiederhole* die obige Abfolge mit einem neuen ankommenden Molekül M′.

Da die Übersetzung der Begriffe unter den entsprechend numerierten Punkten ganz augenscheinlich ist, habe ich mit der Präsentation dieser beiden Listen meinen Beweis geliefert.

Wie können wir uns nun unsere Einsicht zunutze machen, daß die Kybernetik die Wissenschaft des Regelns, Rechnens, Ordnens und der Entropieverzögerung ist? Wir können diese unsere Einsicht natürlich auf jenes System anwenden, das gewöhnlich als die *cause célèbre* allen Regelns, Rechnens, Ordnens und aller Entropieverzögerung angesehen wird, nämlich auf das menschliche Gehirn.

Ich möchte nun nicht den Physikern folgen, die ihre Probleme nach der Anzahl der jeweils untersuchten Objekte ordnen („Das Ein-Körper-Problem", „Das Zwei-Körper-Problem", „Das Drei-Körper-Problem" usw.). Ich werde vielmehr unsere Probleme nach der Anzahl der jeweils betroffenen Gehirne ordnen und im folgenden über das „Ein-Hirn-Problem", das „Zwei-Hirn-Problem", das „Viel-Hirn-" und das „All-Hirn-Problem" sprechen.

1 Das „Ein-Hirn-Problem": die Wissenschaften vom Gehirn

Es liegt auf der Hand, daß die Wissenschaften, die sich mit dem menschlichen Gehirn befassen, eine Theorie des Gehirns, T(G), entwickeln müssen, wenn sie nicht zu einer Physik oder Chemie lebendiger – oder lebendig gewesener – Gewebe degenerieren wollen. Eine solche Theorie muß natürlich von einem Gehirn geschrieben werden: G(T). Daraus folgt, daß eine derartige Theorie so angelegt werden muß, daß sie sich selbst schreibt: T(G(T)).

Eine derartige Theorie wird sich in grundlegender Hinsicht etwa von jener Art der Physik unterscheiden, die sich der (nicht ganz) erfolgreichen Beschreibung einer „subjektlosen Welt" widmet, in der auch der Beobachter selbst keinen Platz haben soll. Ich komme damit zur Verkündung meines Theorems Nr. 3:

> „Die Naturgesetze werden von Menschen geschrieben. Die Gesetze der Biologie müssen sich selbst schreiben."

Man ist versucht, sich zur Widerlegung dieses Theorems auf Gödels Nachweis der Grenzen des Entscheidungsproblems in Systemen zu berufen, die versuchen, über sich selbst zu sprechen. Lars Löfgren und Gotthard Günther haben aber gezeigt, daß Selbsterklärung und Selbstreferenz Begriffe sind, die von Gödels Überlegungen völlig unberührt bleiben. Mit anderen Worten, eine Wissenschaft vom Gehirn im oben skizzierten Sinne ist, so behaupte ich, in der Tat eine völlig legitime Wissenschaft mit einem völlig legitimen Problem.

2 Das „Zwei-Hirn-Problem": Erziehung

Der Großteil unserer institutionalisierten Erziehungsbemühungen hat zum Ziel, unsere Kinder zu trivialisieren. Ich verwende diesen Begriff „Trivialisierung" genau so, wie er in der Automatentheorie gebräuchlich ist. Dort ist eine triviale Maschine durch eine festgelegte Input-Output-Beziehung gekennzeichnet, während in einer nicht-trivialen Maschine (Turingmaschine) der Output durch den Input *und* den internen Zustand der Maschine bestimmt wird. Da unser Erziehungssystem daraufhin angelegt ist, berechenbare Staatsbürger zu erzeugen, besteht sein Zweck darin, alle jene ärgerlichen inneren Zustände auszuschalten, die Unberechenbarkeit und Kreativität ermöglichen. Dies zeigt sich am deutlichsten

in unserer Methode des Prüfens, die nur Fragen zuläßt, auf die die Antworten bereits bekannt (oder definiert) sind, und die folglich vom Schüler auswendiggelernt werden müssen. Ich möchte diese Fragen als „illegitime Fragen" bezeichnen.

Wäre es dagegen nicht faszinierend, sich ein Erziehungssystem vorzustellen, das die zu Erziehenden ent-trivialisiert, indem es sie lehrt, „legitime Fragen" zu stellen, d.h. Fragen, deren Antworten noch unbekannt sind?

3 Das „Viel-Hirn-Problem": Gesellschaft

Es liegt auf der Hand, daß unsere Gesellschaft als ganze von gravierenden Funktionsstörungen befallen ist. Dies äußert sich auf der Ebene des Individuums schmerzhaft in Apathie, Mißtrauen, Gewalt, Isolierung, Ohnmacht, Entfremdung usw. Ich spreche hier von einer „Partizipationskrise", denn das Individuum wird von der Mitwirkung am sozialen Prozeß zunehmend ausgeschlossen. Die Gesellschaft wird zum „System", zum „Establishment" oder was auch immer, zu einem unpersönlichen kafkaesken Monster von eigensinniger Böswilligkeit.

Es fällt nicht schwer zu erkennen, daß der wesentliche Grund für diese Funktionsstörungen im Fehlen des adäquaten Inputs für das Individuum liegt, mit seiner Gesellschaft zu interagieren. Die sogenannten „Kommunikationskanäle", die „Massenmedien", bieten nur eine Einbahnstraße: Sie reden, niemand aber kann darauf antworten. Da der Rückkopplungskanal fehlt, ist das System nicht zu kontrollieren. Die Kybernetik könnte hierfür eine universal zugängliche Apparatur für sozialen Input entwickeln.

4 Das "All-Hirn-Problem": Menschheit

Das bedrückendste Charakteristikum des globalen Systems „Menschheit" ist seine nachweisliche Instabilität und der daraus folgende, unerwartet schnell herannahende Kollaps. Solange die Menschheit sich selbst als ein offenes System behandelt und die Signale der Sensoren ignoriert, die seinen eigenen Zustand vermitteln, bewegen wir uns unaufhaltsam diesem Ende zu. (In der letzten Zeit habe ich mich zu fragen begonnen, ob die Information über den eigenen Zustand die Elemente des Systems überhaupt so rechtzeitig erreichen kann, daß sie noch reagieren können, wenn sie sich entscheiden, zuzuhören statt aufeinander einzuschlagen.)

Das Ziel ist klar: Wir müssen das System schließen, um eine stabile Bevölkerung, eine stabile Wirtschaft und stabile Rohstoffe zu erreichen. Während nun das Problem der Konstruktion eines Kontrollmechanismus für die Bevölkerung und die Wirtschaft mit den geistigen Reserven dieses Planeten gelöst werden kann, müssen wir uns zur Stabilisierung unserer materiellen Ressourcen aufgrund des Zweiten Hauptsatzes der Thermodynamik um außerplanetarische Rohstoffquellen bemühen. Wir verfügen über etwa $2 \cdot 10^{14}$ Kilowatt Sonnenstrahlung. Würde diese in kluger Weise genutzt, könnten die hochstrukturierten und unschätzbaren organischen Ressourcen der Erde, die fossilen wie die lebendigen, für unzählige weitere Generationen gesichert werden.

Wenn wir Ruhm und Erfolg nachjagen, können wir die Tiefe dieser Probleme des Rechnens, des Ordnens, des Regelns und der Entropieverzögerung ignorieren. Da wir als Kybernetiker jedoch angeblich über das Wissen verfügen, sie zu lösen, sollten wir unser Ziel *über* Ruhm und Erfolg setzen und uns still an die Lösung dieser Probleme machen. Wenn wir nämlich unsere Glaubwürdigkeit als Wissenschaftler erhalten wollen, dann kann der erste Schritt nur darin bestehen, unser Wissen auf uns selbst anzuwenden und eine Weltgesellschaft zu bilden, die nicht so sehr *für* die Kybernetik da ist, sondern vielmehr kybernetisch *funktioniert.* So nämlich verstehe ich Dennis Gabors Aufruf in einer früheren Nummer dieser Zeitschrift: „Kybernetiker dieser Welt, vereinigt euch!" Ohne Kommunikation gibt es keine Regelung; ohne Regelung gibt es kein Ziel; und ohne ein Ziel werden Begriffe wie „Gesellschaft" oder „System" zu Leerformeln.

Wissen bedeutet Verantwortung. Ein Arzt muß direkt an der Unfallstelle tätig werden. Wir können es uns nicht länger leisten, einer globalen Katastrophe lediglich als wissende Zuschauer zuzusehen. Wir müssen all das Wissen, das wir haben, durch Kommunikation und Kooperation miteinander teilen und damit die Probleme unserer Zeit bewältigen. Nur auf diese Weise können wir unsere soziale und individuelle Verantwortung als Kybernetiker erfüllen, nur indem wir das praktizieren, was wir predigen.

Über das Konstruieren von Wirklichkeiten *

Kurzfassung: „Triff eine Unterscheidung!"
(Brown 1972,3)

Das Postulat

Sicher kennen Sie den Bürger Jourdain in Molières Stück „Der Bürger als Edelmann", der sich als Neureicher in den gebildeten Kreisen der französischen Aristokratie bewegt und großen Lerneifer an den Tag legt. Als eines Tages unter seinen neuen Freunden von Dichtung und Prosa die Rede ist, entdeckt Jourdain zu seinem Erstaunen und mit übergroßer Freude, daß er immer dann, wenn er spricht, *Prosa* spricht. Er ist von dieser Entdeckung überwältigt: „Ich spreche Prosa! Ich habe immer schon Prosa gesprochen! Ich habe mein ganzes Leben lang Prosa gesprochen!"

Eine ähnliche Entdeckung ist vor nicht allzu langer Zeit gemacht worden, es handelte sich dabei aber weder um Dichtung noch um Prosa, – es war die Umwelt, die entdeckt wurde. Ich erinnere mich noch, wie vor vielleicht 10 oder 15 Jahren einige meiner amerikanischen Freunde zu mir gelaufen kamen, voll der Freude und des Erstaunens über ihre gerade gemachte große Entdeckung: „Ich lebe in einer Umwelt! Ich habe immer schon in einer Umwelt gelebt! Ich habe mein ganzes Leben lang in einer Umwelt gelebt!"

Weder Monsieur Jourdain noch meine Freunde haben aber bis heute eine andere Entdeckung gemacht: Wenn Monsieur Jourdain spricht, sei es Prosa oder Dichtung, dann ist er selbst es, der diese erfindet; und immer dann, wenn wir unsere Umwelt wahrnehmen, sind wir selbst es, die diese Umwelt erfinden.

Jede Entdeckung hat ihre leidvolle und ihre erfreuliche Seite: Es bereitet Schmerzen, mit einer neuen Einsicht fertig zu werden, und Freude, diese Einsicht gewonnen zu haben. Ich sehe den einzigen Zweck meines Vortrags darin, für all jene, die diese Entdeckung noch nicht gemacht haben, die Leiden zu minimieren und die Freude zu maximieren, andererseits all jene, die diese Entdeckung bereits gemacht haben, wissen zu lassen, daß sie nicht allein sind. Die Entdeckung, die jeder von uns für sich selbst machen muß, ist im folgenden Postulat ausgedrückt:

Die Umwelt, die wir wahrnehmen, ist unsere Erfindung.

Es liegt nun an mir, diese unerhörte Behauptung zu rechtfertigen. Ich will dies folgendermaßen tun: Zunächst möchte ich Sie einladen, an einem Experiment teilzunehmen; darauf werde ich einen klinischen Fall sowie die Ergebnisse zweier Experimente darstellen. Danach möchte ich eine Interpretation und sodann eine

---

* Gekürzte Fassung eines Vortrags vom 15. April 1973 zur Eröffnung der "Fourth International Conference on Environmental Design Research" am Virginia Polytechnic Institute, Blacksburg/Virginia.

stark komprimierte Darstellung der neurophysiologischen Basis dieser Experimente und meines eben formulierten Postulats anbieten. Abschließend möchte ich versuchen, die Bedeutung alles dieses für ästhetische und ethische Überlegungen klarzumachen.

Die Experimente

Bild 1

1 Der blinde Fleck

Halten Sie das Buch mit der rechten Hand, schließen Sie das linke Auge und fixieren Sie den Stern in Bild 1 mit dem rechten Auge. Bewegen Sie sodann das Buch langsam entlang der Sehachse vor und zurück, bis der Abstand erreicht ist (ca. 30 bis 35 cm), bei dem der runde schwarze Fleck verschwindet. Wenn der Stern gut fixiert wird, bleibt der Fleck unsichtbar, auch wenn das Buch langsam parallel zu sich selbst in beliebiger Richtung bewegt wird. Diese lokalisierte Blindheit ist eine direkte Folge des Fehlens von Photorezeptoren (Stäbchen und Zapfen) an dem Punkt der Retina, der „Scheibe", wo alle Fasern von der lichtempfindlichen Schicht des Auges zusammenkommen und den Sehnerv bilden. Klarerweise kann der schwarze Fleck, wenn er auf den blinden Fleck projiziert wird, nicht gesehen werden. Es ist zu betonen, daß diese lokalisierte Blindheit nicht als schwarzer Fleck in unserem visuellen Feld wahrgenommen wird (einen schwarzen Fleck sehen würde bedeuten, daß man „sieht"), sondern daß diese Blindheit *überhaupt* nicht wahrgenommen wird, d. h. weder als etwas das gegeben ist, noch als etwas, das fehlt: Wir sehen nicht, daß wir nicht sehen.

2 Skotom

Bestimmte gut lokalisierte Okzipitalläsionen im Gehirn (z.B. Verletzungen durch Hochgeschwindigkeitsprojektile) heilen relativ schnell, ohne daß sich der Patient irgendeines bemerkbaren Verlustes seines Sehvermögens bewußt wird. Nach einigen Wochen macht sich jedoch beim Patienten eine gewisse motorische Dysfunktion bemerkbar, z. B. der Verlust der Kontrolle von Arm- oder Beinbewegungen auf der einen oder der anderen Seite und ähnliches. Klinische Tests zeigen jedoch, daß im motorischen System nichts fehlt, daß es aber in einigen Fällen zu einem beträchtlichen Verlust eines großen Teils des visuellen Feldes

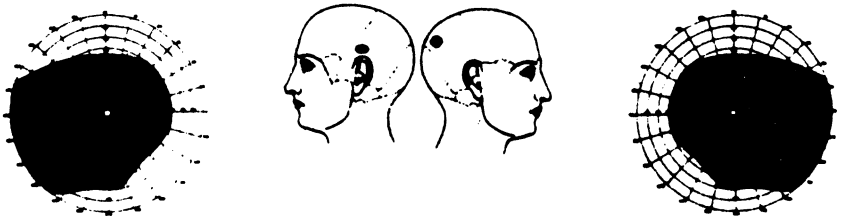


Bild 2

kommt (*Skotom*; vgl. Teuber 1961). Die erfolgreiche Therapie besteht darin, dem Patienten über einen Zeitraum von ein bis zwei Monaten die Augen zu verbinden, bis er die Beherrschung seines motorischen Systems dadurch wiedergewinnt, daß er seine „Aufmerksamkeit" von (nicht-existenten) visuellen Hinweisen auf seine Körperstellung auf jene (normal arbeitenden) Kanäle umstellt, die *direkte* Hinweise auf seine Körperstellung aus (propriozeptiven) Sensoren in Muskeln und Gelenken liefern. Wiederum ist zu betonen, daß das „Fehlen von Wahrnehmung" *nicht* wahrgenommen und daß die Fähigkeit wahrzunehmen durch sensumotorische Interaktion wieder aufgebaut wird. Dies läßt mich zwei Metaphern formulieren:

(a) „Wahrnehmen ist Handeln".

(b) „Wenn ich nicht sehe, daß ich blind bin, dann bin ich blind; wenn ich aber sehe, daß ich blind bin, dann sehe ich".

3 Alternanten

Ein einzelnes Wort wird auf ein Tonband gesprochen und das Band wird sodann (ohne Knackgeräusch) zu einer Schleife zusammengeklebt. Das Wort wird daraufhin immer wieder mit eher großer als kleiner Lautstärke abgespielt. Nach ein bis zwei Minuten des Zuhörens (nach 50 bis 150 Wiederholungen) verwandelt sich das bis dahin klar wahrgenommene Wort ganz abrupt in ein anderes bedeutsames und klar wahrgenommenes Wort: in eine „Alternante". Nach 10 bis 30 Wiederholungen dieser ersten Alternante springt diese plötzlich in eine zweite Alternante um usw. (Naeser/Lilly 1971). Ich gebe im folgenden eine kleine Auswahl der 758 Alternanten, die von etwa 200 Versuchspersonen genannt wurden, denen das Wort COGITATE in der geschilderten Weise immer wieder vorgespielt wurde: AGITATE; ANNOTATE; ARBITRATE; ARTISTRY; BACK AND FORTH; BREVITY; CA D'ETAIT; CANDIDATE; CAN'T YOU SEE; CAN'T YOU STAY; CAPE COD YOU SAY; CARD ESTATE; CARDIO TAPE; CAR DISTRICT; CATCH A TAPE; CAVITATE; CHA CHA CHE; COGITATE; COMPUTATE; CONJUGATE; CONSCIOUS STATE; COUNTER TAPE; COUNT TO TEN; COUNT TO THREE; COUNT YER TAPE; CUT THE STEAK; ENTITY; FANTASY; GOD TO TAKE; GOD YOU SAY; GOT A DATE; GOT YOUR PAY; GOT YOUR TAPE; GRATITUDE; GRAVITY; GUARD THE TIT; GURGITATE; HAD TO TAKE; KINDS OF TAPE; MAJESTY; MARMALADE ....

4 Verstehen (urspr. „wahrnehmen, auffassen, begreifen, etwas können")

An verschiedenen Stellen der Hörbahn im Gehirn einer Katze werden Mikroelektroden so eingesetzt, daß Ableitungen (Elektroenzephalogramme) sowohl von jenen Nervenzellen möglich sind, die als erste auditorische Stimuli empfangen (Cochlea Nucleus, CN), als auch von anderen bis hin zum auditorischen Cortex (Worden 1959). Die derart präparierte Katze wird in einen Käfig gesetzt, in dem sich ein Futterbehälter befindet, dessen Deckel durch Niederdrücken eines Hebels geöffnet werden kann. Die Verbindung zwischen Hebel und Deckel funktioniert jedoch nur dann, wenn wiederholt ein bestimmter Einzelton (in diesem Falle C_6, d.h. etwa 1000 Hz) präsentiert wird. Die Katze muß lernen, daß C_6 Nahrung „bedeutet". Die Bilder 3 bis 6 zeigen die Muster der Nerventätigkeit an acht Stellen der Hörbahn in aufsteigender Ordnung und bei vier aufeinanderfolgenden Stadien dieses Lernprozesses (Worden 1959). Das Verhalten der Katze, das mit

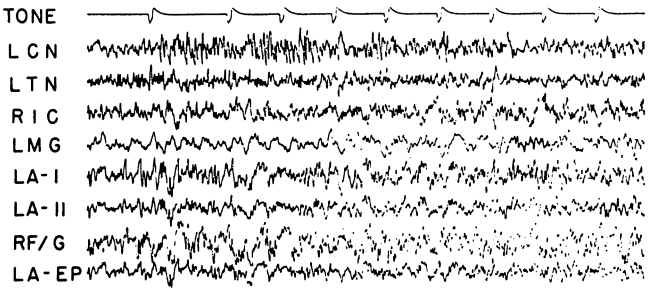


Bild 3

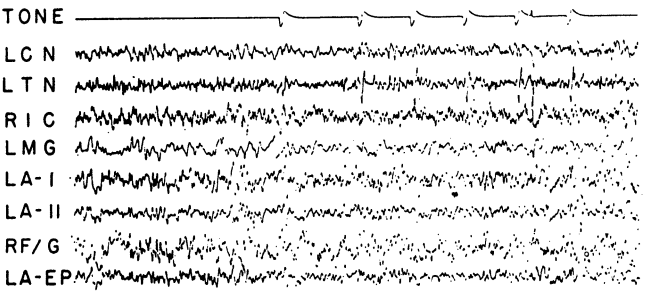


Bild 4

Bild 5

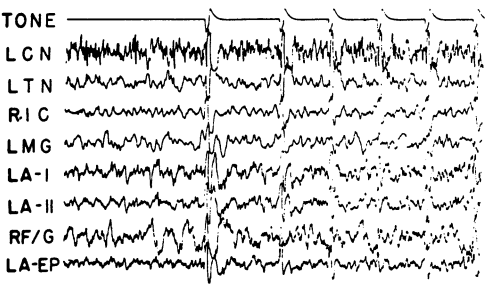


Bild 6

den abgeleiteten neuronalen Aktivitäten verknüpft ist, ist für Bild 3 „zufällige Suche", für Bild 4 „Prüfung des Hebels", für Bild 5 „Hebel sofort niedergedrückt", und für Bild 6 „Katze geht direkt auf den Hebel zu (volles Verstehen)". Es ist zu betonen, daß kein Ton wahrgenommen wird, solange dieser Ton uninterpretierbar ist (Bilder 3 und 4: reines Geräusch), daß das ganze System aber sofort zu arbeiten beginnt, wenn der erste „Piepton" zu hören ist (Bilder 5 und 6: Geräusch wird zum Signal), wenn also eine Wahrnehmung verständlich wird, wenn *unsere* Wahrnehmung von „piep", „piep", „piep" in der Wahrnehmung der *Katze* „Nahrung", „Nahrung", „Nahrung" bedeutet.

Interpretation

Ich habe in diesen Experimenten Beispiele zitiert, bei denen wir etwas sehen oder hören, was gar nicht „da" ist, oder bei denen wir *nicht* sehen oder hören, was „da" ist, es sei denn, unsere Koordination von Sinneswahrnehmung und Bewegung erlaubt uns, das, was da zu sein scheint, zu „erfassen". Ich möchte diese Beobachtung noch dadurch bekräftigen, daß ich mein „Prinzip der undifferenzierten Kodierung" vortrage:

> „Die Reaktion einer Nervenzelle enkodiert *nicht* die physikalischen Merkmale des Agens, das ihre Reaktion verursacht. Es wird lediglich das ‚so viel' an diesem Punkt meines Körpers enkodiert, nicht aber das ‚was'."

Betrachten wir z.B. eine lichtempfindliche Rezeptorzelle in der Retina, ein „Stäbchen", das die von einer fernen Quelle ausgehende elektromagnetische Strahlung absorbiert. Diese Absorption verursacht eine Veränderung des elektrochemischen Potentials des Stäbchens, die schließlich eine periodische elektrische Entladung in Zellen auf einer höheren Ebene des postretinalen Netzwerks verursacht (s.u.), und zwar mit einer Periodizität, die der Intensität der absorbierten Strahlung proportional ist, die aber kein Anzeichen dafür enthält, daß es elektromagnetische Strahlung war, die das Stäbchen zu feuern veranlaßte. Das gleiche gilt für jeden beliebigen anderen sensorischen Rezeptor, die Geschmacksknöspchen, die Druckrezeptoren und alle die anderen Rezeptoren, die mit den Sinneswahrnehmungen des Geruchs, der Wärme und Kälte, der Klänge und Geräusche usw. verbunden sind: Sie sind alle „blind", was die *Qualität* ihrer Stimulierung angeht und reagieren nur auf deren *Quantität*. Auch wenn wir dies überraschend finden, sollte es uns doch nicht verwundern: „da draußen" gibt es nämlich in der Tat weder Licht noch Farben, sondern lediglich elektromagnetische Wellen; „da draußen" gibt es weder Klänge noch Musik, sondern lediglich periodische Druckwellen der Luft; „da draußen" gibt es keine Wärme und keine Kälte, sondern nur bewegte Moleküle mit größerer oder geringerer durchschnittlicher kinetischer Energie usw. Und schließlich gibt es „da draußen" sicherlich keinen Schmerz. Da nun die physikalischen Eigenschaften des Stimulus – seine *Qualität* – von der Nervenaktivität nicht enkodiert werden, stellt sich die fundamentale Frage, wie unser Gehirn denn die überwältigende Vielfalt dieser farbenprächtigen Welt hervorzaubern kann, wie wir sie in jedem Augenblick unseres bewußten Lebens erfahren, – und manchmal sogar, wenn wir schlafen und träumen. Dies ist das „Problem der Kognition", die Suche nach einer Erklärung der kognitiven Prozesse. Die Art, in der eine Frage gestellt wird, bestimmt die Art, in der sie beantwortet werden kann. Es liegt also nun an mir, das „Problem der Kognition" auf solche Weise zu formulieren, daß

die theoretischen Werkzeuge eingesetzt werden können, über die wir heute verfügen. Zu diesem Zwecke möchte ich „Kognition" auf folgende Weise umschreiben (→):

KOGNITION → Errechnung einer Realität.

Damit provoziere ich sicherlich einen Sturm entrüsteter Einwände. Zunächst scheint es, als ob ich den unbekannten Begriff „Kognition" durch drei andere Begriffe ersetze, wovon zwei, „Errechnung" und „Realität", noch weitaus obskurer sind als mein Definiendum, während das einzige bestimmte Wort, das ich gebrauche, der unbestimmte Artikel „eine" ist. Darüber hinaus legt der Gebrauch des unbestimmten Artikels die lächerliche Vorstellung nahe, daß es noch andere Realitäten neben „der" einen und einzigen Realität, nämlich unserer geliebten eigenen Umwelt gibt; und schließlich scheint es, als ob ich durch den Gebrauch des Ausdrucks „Errechnung" ausdrücken möchte, daß alles, von meiner Armbanduhr bis zu den Milchstraßensystemen des Universums, bloß errechnet wird und nicht schlicht „da" ist. Unerhört!

Lassen sie mich diese Einwände einen nach dem anderen abhandeln. Zunächst möchte ich das semantische Unbehagen beseitigen, das Ausdrücke wie „Rechnen" oder „Errechnung" in einer Versammlung von Damen und Herren verursachen müssen, die stärker den Geisteswissenschaften als den Naturwissenschaften zuneigen.

Das Wort „rechnen" kommt von einem im Hochdeutschen nicht mehr vorhandenen Adjektiv, das „ordentlich, genau" bedeutet. „Rechnen" heißt also ursprünglich „in Ordnung bringen, ordnen". Dazu gehört u.a. auch „Rechenschaft" und „recht". Es braucht damit also keineswegs auf numerische Größen Bezug genommen werden.

Ich möchte den Begriff des „Rechnens" in diesem sehr allgemeinen Sinn verwenden, um jede (nicht notwendig numerische) *Operation* zu benennen, die beobachtete physikalische Entitäten („Objekte") transformiert, modifiziert, ordnet, neu anordnet usw. So werde ich z.B. die einfache Permutation der drei Buchstaben A,B,C zu C,A,B, zu einer Anordnung also, in der der letzte Buchstabe an die erste Stelle rückt, als „Rechnen" bezeichnen. In ähnlicher Weise nenne ich die Operation eine Errechnung, die die Kommas zwischen den Buchstaben beseitigt: CAB, in gleicher Weise die semantische Transformation, die CAB zu TAXI verändert usw. Als nächstes möchte ich den Gebrauch des unbestimmten Artikels im Ausdruck „eine Realität" verteidigen. Ich könnte mich hier natürlich hinter dem logischen Argument verschanzen, daß ich mit meiner Lösung des allgemeinen Falles, wie er durch das „ein" angedeutet wird, auch jeden speziellen Fall, wie er durch den Gebrauch des „die" ausgedrückt wird, gelöst habe. Meine Absicht ist aber viel radikaler. Es gibt eine tiefe Kluft zwischen dem „die"-Denken und dem „eine"-Denken, wofür wiederum die Begriffe „Bestätigung" bzw. „Korrelation" jeweils als erklärende Paradigmen der Wahrnehmung gelten. Die „die"-Auffassung meint: Meine Wahrnehmung der Berührung ist eine *Bestätigung* meiner visuellen Wahrnehmung, daß es einen Tisch *gibt*. Die „eine"-Auffassung meint: Meine Wahrnehmung der Berührung in *Korrelation* mit meiner visuellen Sinneswahrnehmung *erzeugt* eine Erfahrung, die ich als „Hier ist ein Tisch" beschreiben kann. Ich lehne die „die"-Position aus epistemologischen Gründen ab, denn auf diese Weise wird das ganze Problem der Kognition einfach in den blinden Fleck des Erkennens abgedrängt: Man merkt nicht einmal mehr, daß man das Problem der Kognition nicht sieht.

Schließlich könnte man durchaus zu Recht darauf hinweisen, daß kognitive Prozesse weder Armbanduhren noch Milchstraßensysteme errechnen, sondern im besten Fall *Beschreibungen* derartiger Entitäten. Diesem Einwand will ich stattgeben, und ich ersetze daher meine frühere Umschreibung durch die folgende:

KOGNITION → Errechnung von Beschreibungen einer Realität.

Neurophysiologen sagen uns jedoch (Maturana 1970a), daß eine auf einer bestimmten Ebene neuronaler Aktivität errechnete Beschreibung, etwa ein auf die Retina projiziertes Bild, auf einer höheren Ebene erneut bearbeitet wird, danach wieder usw., wobei bestimmte motorische Aktivitäten von einem Beobachter als „terminale Beschreibungen" angesehen werden können, wie z.B. die folgende Äußerung: „Hier ist ein Tisch". Ich habe daher meine Umschreibung erneut zu modifizieren, so daß sie nun folgendermaßen lautet:

KOGNITION → Errechnung von Beschreibungen ⟲

wobei der zurückführende Pfeil ausdrücken soll, daß es sich hier um eine infinite Rekursion von Beschreibungen von Beschreibungen.. usw. handelt.

Diese Formulierung hat den Vorteil, daß eine Unbekannte, nämlich „Realität", mit Erfolg ausgeschaltet worden ist. Realität erscheint nur mehr implizit als die Aktivität rekursiver Beschreibungen. Schließlich können wir auf die Tatsache zurückgreifen, daß die Errechnung von Beschreibungen natürlich nichts anderes ist als eine Errechnung. Es ergibt sich somit:

KOGNITION → Errechnungen von ⟲

Ich fasse zusammen: Mein Vorschlag besteht darin, kognitive Prozesse als nie endende rekursive Prozesse des (Er-)Rechnens aufzufassen. Ich hoffe, daß ich nun mit der folgenden *tour de force* durch die Neurophysiologie diese Interpretation verständlich machen kann.

Neurophysiologie

1 Evolution

Um volles Verständnis dafür zu erreichen, daß das Prinzip der rekursiven Errechnung in der Tat allen kognitiven Prozessen zugrunde liegt – ja sogar dem Leben schlechthin, wie mir einer der fortgeschrittensten Denker der Biologie versichert (Maturana 1970b) –, kann es hilfreich sein, für einen Augenblick auf die elementarsten – oder wie die Evolutionstheoretiker sagen würden, auf sehr „frühe" – Manifestationen dieses Prinzips zurückzugehen. Es handelt sich dabei um die „unabhängigen Effektoren" bzw. unabhängigen sensumotorischen Einheiten, wie sie sich über die Oberflächen von Einzellern und Vielzellern verteilt finden (Bild 7).

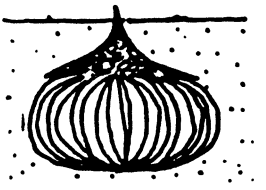


Bild 7

Der dreieckige Kopf dieser Einheit, der mit seiner Spitze aus der Oberfläche hervorwächst, ist der sensorische Teil, der zwiebelartige Körper ist der kontraktile motorische Teil. Eine Veränderung der chemischen Konzentration eines Agens in der unmittelbaren Nachbarschaft der sensiblen Spitze, für die das Agens „wahrnehmbar" ist, verursacht die sofortige Kontraktion der Einheit. Die daraus resultierende Bewegung durch eine Veränderung der Gestalt des Lebewesens oder seiner örtlichen Lage kann ihrerseits wahrnehmbare Veränderungen der Konzentration des Agens in der Nachbarschaft solcher Einheiten erzeugen, die wiederum deren sofortige Kontraktion verursachen... usw. Es ergibt sich daher die folgende Rekursion:

→ Veränderung der Sinneswahrnehmung → Veränderung der Gestalt →

Die Trennung der Orte der Sinneswahrnehmung von denen der Handlung scheint der nächste evolutionäre Schritt gewesen zu sein (Bild 8).

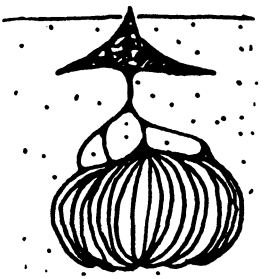


Bild 8

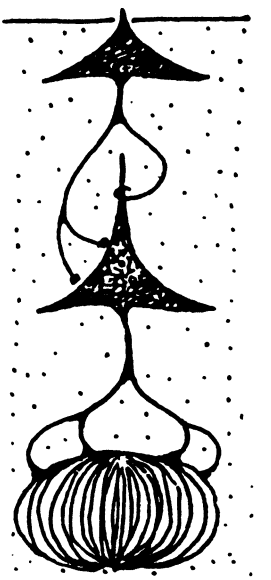


Bild 9

Die sensorischen und die motorischen Organe sind nun durch dünne Fasern miteinander verbunden, die „Axone" (im allgemeinen degenerierte Muskelfasern, die ihre Kontraktilität verloren haben), die die Einwirkungen auf die Sensoren an die zugehörigen Effektoren übermitteln und somit den Begriff des „Signals" entstehen lassen: *hier* wird gesehen, *dort* wird entsprechend gehandelt. Der entscheidende Schritt in der Evolution der komplexen Organisation des Zentralnervensystems (ZNS) der Säugetiere scheint jedoch das Auftreten eines „internuntialen Neurons" gewesen zu sein, einer Zelle, die zwischen der sensorischen und der motorischen Einheit gelagert ist (Bild 9).

Es handelt sich dabei im allgemeinen um eine sensorische Zelle, die jedoch so spezialisiert ist, daß sie nur auf ein universales „Agens" reagiert, nämlich auf die elektrische Aktivität der afferenten Axone, die in ihrer Nachbarschaft enden. Da

ihre gegenwärtige Aktivität ihre spätere Reaktionsfähigkeit beeinflussen kann, führt diese Zelle das Element des Rechnens in das Reich der Lebewesen ein und ermöglicht so den entsprechenden Organismen die erstaunliche Vielfalt nichttrivialer Verhaltensweisen. Ist einmal der genetische Kode zur Herstellung des internuntialen Neurons entwickelt, dann bedeutet es in der Tat nur mehr geringen Aufwand, den genetischen Befehl „Wiederholen" hinzuzufügen. Ich glaube daher, daß nun leicht einzusehen ist, warum sich diese Neurone entlang zusätzlicher vertikaler Schichten mit zunehmenden horizontalen Verbindungen so schnell vermehrt haben, um jene komplexen ineinander verknüpften Strukturen zu bilden, die wir „Gehirne" nennen.

2 Das Neuron

Unser Gehirn besteht aus mehr als zehn Milliarden „Neuronen", höchst spezialisierten Einzelzellen mit drei anatomisch verschiedenen Merkmalen (Bild 10):

(a) den „Dendriten", das sind zweigähnliche Verästelungen, die sich nach oben und seitwärts ausdehnen;
(b) dem „Zelleib", einer Knolle im Zentrum, die den Zellkern enthält; und
(c) dem „Axon", der glatten Faser, die sich nach unten erstreckt.

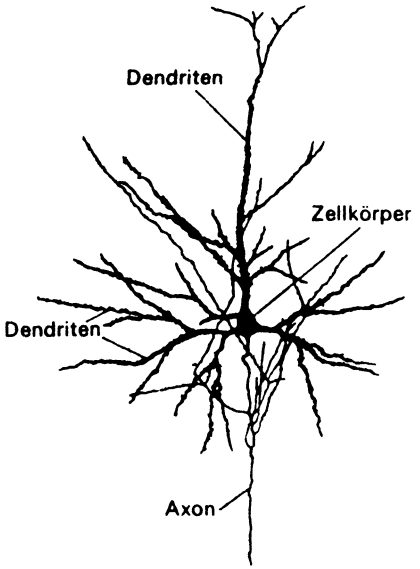


Bild 10

Die einzelnen Ausläufer des Axons enden auf den Dendriten eines anderen (manchmal aber auch (rekursiv) desselben) Neurons. Die Membran, die den Zellleib einhüllt, bildet auch die röhrenartige Scheide für die Dendriten und Axone und ist verantwortlich dafür, daß das Zellinnere gegenüber seiner Umgebung eine elektrische Ladung von etwa einem Zehntel Volt aufweist. Wird diese Ladung in der Region der Dendriten hinreichend gestört, dann „feuert" das Neuron und schickt diese Störung entlang seiner Axone an dessen Endpunkte, die Synapsen.

3 Die Übertragung

Da derartige Störeinwirkungen elektrischer Art sind, können sie von Mikrosonden aufgenommen, verstärkt und aufgezeichnet werden. Bild 11 zeigt drei Beispiele periodischer Entladungen eines Druckrezeptors, der ständig stimuliert wird, wobei die niedrigere Frequenz einem schwachen, die hohe Frequenz einem starken Stimulus entspricht. Die Größe der Entladung ist klarerweise überall dieselbe, wobei die Impulsfrequenz die Stimulusintensität und nur diese repräsentiert.

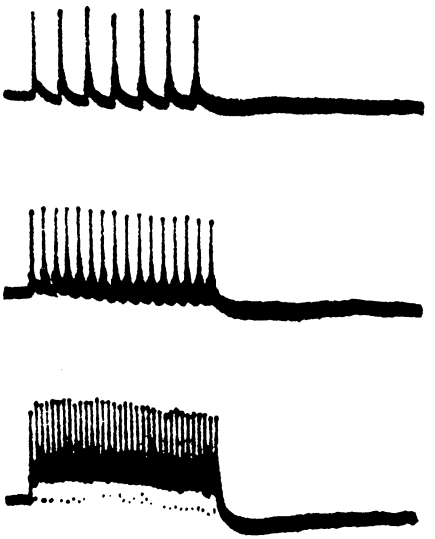


Bild 11

4 Die Synapse

Bild 12 zeigt eine synaptische Verbindung.

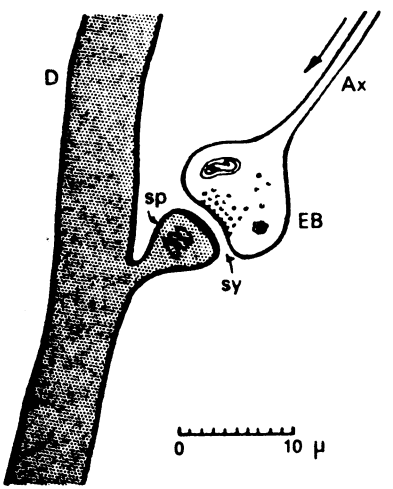


Bild 12

Das afferente Axon (Ax), entlang dessen die Impulse sich ausbreiten, endet in einer Endknolle (EB), die vom Ast (sp) eines Dendriten (D) des Zielneurons durch einen winzigen Spalt (sy), den „synaptischen Spalt", getrennt ist. (Die zahlreichen Äste dieser Art bewirken das stachelige Aussehen der Dendriten in Bild 10.) Die chemische Zusammensetzung der „Transmittersubstanzen", die den synaptischen Spalt ausfüllen, ist entscheidend für die Wirkung, die ein ankommender Impuls auf das Neuron ausüben kann: Unter bestimmten Umständen kann der Impuls einen „inhibitorischen Effekt" haben (Aufhebung eines anderen gleichzeitig ankommenden Impulses), unter anderen einen „Bahnungseffekt" (Verstärkung eines anderen Impulses, der das Neuron aktiviert). Der synaptische Spalt kann folglich als die „Mikroumwelt" einer sensiblen Spitze, nämlich des Dendritenasts, angesehen werden, und mit dieser Interpretation vor Augen können wir die Sensitivität des Zentralnervensystems gegenüber Veränderungen der *inneren* Umwelt (der Gesamtsumme aller Mikroumwelten) mit seiner Sensitivität gegenüber Veränderungen der *äußeren* Umwelt (das heißt aller sensorischen Rezeptoren) vergleichen. Da es lediglich einige 100 Millionen sensorische Rezeptoren und etwa 10 000 Milliarden Synapsen in unserem Nervensystem gibt, sind wir gegenüber Veränderungen in unserer inneren Umwelt 100 000mal stärker empfindlich als gegenüber Veränderungen in unserer äußeren Umwelt.

5 *Der Cortex*

Um nun zumindest eine gewisse Vorstellung von der Organisation der gesamten Maschinerie zu gewinnen, die alle unsere perzeptuellen, intellektuellen und emotionalen Erfahrungen errechnet, habe ich Bild 13 (Sholl 1956) beigegeben, das einen vergrößerten Schnitt von etwa zwei Quadratmillimetern aus dem Cortex einer Katze zeigt, der mit Hilfe einer besonderen Technik so gefärbt worden ist, daß lediglich die Zellkörper und Dendriten von etwa 1 % aller vorhandenen Neuronen erkennbar sind. Auch wenn man sich die vielen Querverbindungen zwischen diesen Neuronen, wie sie durch die (hier unsichtbaren) Axone hergestellt werden,

Bild 13

ebenso bloß vorstellen muß wie die Packungsdichte, die hundertmal größer ist als die gezeigte, dürfte sich daraus die Rechenmächtigkeit allein dieses in der Tat äußerst kleinen Teils des Gehirns erahnen lassen.

6 *Descartes*

Dieses Bild ist natürlich weit entfernt von dem, das man sich vor etwa dreihundert Jahren gemacht hat (Descartes 1664): „Wenn das Feuer A dem Fuß B nahekommt (Bild 14), dann haben die Teilchen dieses Feuers, die sich, wie wir wissen, mit großer Geschwindigkeit bewegen, die Kraft, jenen Teil der Haut des Fußes, den sie berühren, zu bewegen; auf diese Weise ziehen sie an dem dünnen Faden c, den wir an den Wurzeln der Zehen und an den Nerven angebunden sehen; gleichzeitig öffnen sie den Zugang zu der Pore d und e, wo dieser dünne Faden endet, so wie das Ziehen an dem einen Ende einer Kordel die Glocke läuten läßt, die an deren anderem Ende hängt. Das nun so bewirkte Öffnen der Pore bzw. des kleinen Ausgangs d und e läßt die Lebensgeister des Hohlraums F austreten und fortströmen, zum Teil in Muskeln, die den Fuß vom Feuer zurückziehen helfen, zum Teil in andere, die die Augen und den Kopf drehen, um das Feuer anzusehen, und zum Teil in wieder andere, die die Hände darauf hinbewegen und den ganzen Körper neigen, um den Fuß zu schützen."

Man bedenke, daß einige unserer heutigen Behavioristen immer noch diese Auffassung vertreten (Skinner 1971), lediglich mit dem Unterschied, daß Descartes' „Lebensgeister" inzwischen verlorengegangen sind.

Bild 14

7 *Die Errechnung*

Die Retina der Wirbeltiere und das damit verbundene Nervengewebe sind ein typischer Fall neuronaler Errechnung. Bild 15 bietet eine schematische Darstellung einer Säugetierretina und ihres postretinalen Netzwerks.

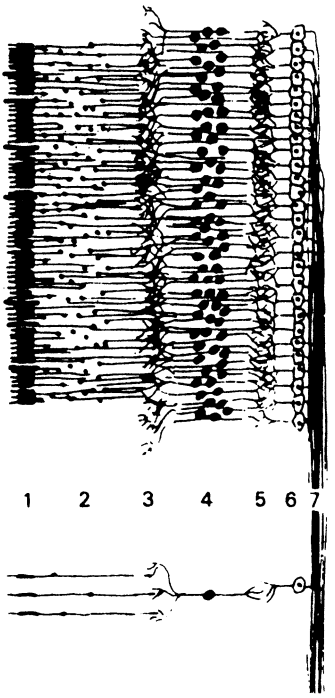


Bild 15

Die mit # 1 bezeichnete Schicht zeigt die Anordnung von Stäbchen und Zapfen, Schicht # 2 die Körper und Kerne dieser Zellen. Schicht # 3 zeigt die Region, in der die Axone der Rezeptoren mit den dendritischen Verzweigungen der „Bipolarzellen" Synapsen bilden, die ihrerseits in der Schicht # 5 mit den Dendriten der „Ganglienzellen" (# 6) Synapsen bilden, deren Aktivität in noch tieferliegende Schichten des Gehirns über jene Axone übermittelt wird, die zum Sehnerv (# 7) gebündelt sind. Rechenprozesse finden innerhalb der beiden Schichten # 3 und # 5 statt, das heißt dort, wo sich die Synapsen befinden. Wie Maturana/Uribe/Frenk (1968) gezeigt haben, werden eben dort die Sinneswahrnehmungen der Farbe sowie einige Formmerkmale errechnet.

Zur Formberechnung: Betrachten wir das zweischichtige periodische Netzwerk in Bild 16, in dem die obere Schicht Rezeptorzellen enthält, die „licht"-

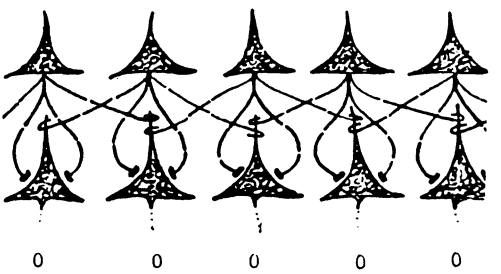


Bild 16

empfindlich sind. Jeder dieser Rezeptoren ist mit drei Neuronen in der darunter liegenden (rechnenden) Schicht verbunden, wobei zwei exzitatorische Synapsen auf dem einen direkt darunter liegenden Neuron enden (symbolisiert durch die dem Zellkörper angefügten Knöpfe), und eine inhibitorische Synapse (symbolisiert durch eine Schleife um die Spitze) die beiden anderen Neurone, eines links und eines rechts, erfaßt. Es ist klar, daß die rechnende Schicht nicht reagieren wird, wenn Licht einheitlich auf die ganze rezeptive Schicht projiziert wird, denn die beiden exzitatorischen Stimuli eines Rechnerneurons werden durch die inhibitorischen Signale, die von den beiden lateralen Rezeptoren kommen, aufgehoben. Diese Nullreaktion wird unter der stärksten wie auch unter der schwächsten Stimulierung ebenso wie unter langsamer oder rascher Veränderung der Belichtung immer gleich erfolgen. Es stellt sich nun die berechtigte Frage „Warum dieser komplexe Apparat, der nichts tut?"

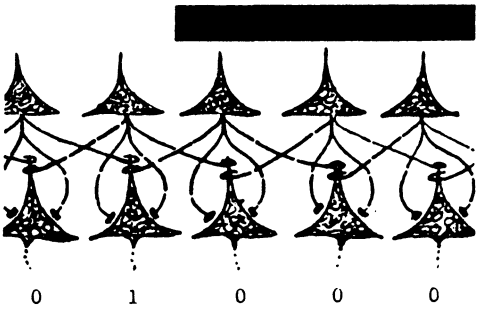


Bild 17

Nehmen wir nun aber Bild 17, in dem der Bahn des Lichts, das auf die Schicht der Rezeptoren fällt, ein Hindernis entgegengesetzt wird. Wiederum werden alle Neuronen der unteren Schicht ruhig bleiben, ausgenommen das eine Neuron am Rande des Hindernisses, denn es erhält zwei exzitatorische Signale vom darüberliegenden Rezeptor, aber nur *ein* inhibitorisches Signal von dem links davon liegenden Sensor. Wir verstehen nun die wichtige Funktion dieses Netzes, denn es errechnet jede räumliche *Veränderung* des visuellen Feldes dieses „Auges", und zwar unabhängig von der Intensität des umgebenden Lichts und seiner zeitlichen Schwankungen, und unabhängig von Ort und Ausdehnung des Hindernisses. Auch wenn alle Operationen, die diese Rechenprozesse ausmachen, elementarer Art sind, erlaubt uns die Organisation dieser Operationen, ein Prinzip von beträchtlicher Tragweite zu erkennen, nämlich jenes der Errechnung von „abstrakten" Vorstellungen, in diesem Fall der Vorstellung „Kante". Ich hoffe, daß dieses einfache Beispiel ausreicht, um die Möglichkeit der Generalisierung dieses Prinzips zu veranschaulichen, und zwar dahingehend, daß Errechnung sich zumindest auf zwei Ebenen zeigt, nämlich (a) in den tatsächlich ausgeführten Operationen, und (b) in der Organisation dieser Operationen, wie sie hier durch die Struktur des Nervennetzes dargestellt wird. In der Computersprache würde man bei (a) von „Operationen" sprechen, bei (b) jedoch von einem „Programm". Wie wir später noch sehen werden, können in „biologischen Rechnern" die Programme selbst zum Gegenstand von Rechenprozessen werden. Wir erreichen so „Metaprogramme", „Meta-Metaprogramme"... usw. Und all das ist natürlich die Folge der rekursiven Organisation dieser Systeme.

8 *Geschlossenheit*

Vielleicht haben wir nun durch die Konzentration unserer Aufmerksamkeit auf die neurophysiologischen Einzelteile den Organismus als funktionierende Ganzheit aus den Augen verloren. In Bild 18 habe ich daher die Einzelstücke wieder durch ihre funktionalen Beziehungen verbunden.

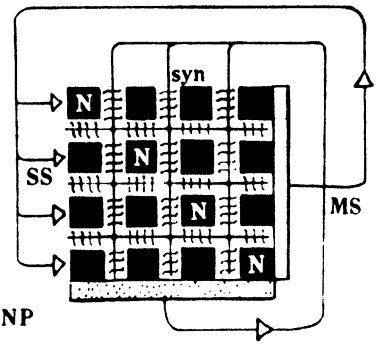


Bild 18

Die mit N bezeichneten schwarzen Quadrate stellen Neuronenbündel dar, die mit Neuronen anderer Bündel über die (synaptischen) Spalte, die durch die Zwischenräume zwischen den Quadraten angezeigt sein sollen, Synapsen bilden. Die sensorische Oberfläche (SS) des Organismus befindet sich auf der linken Seite, seine motorische Oberfläche (MS) auf der rechten, und die Hirnanhangdrüse (Hypophyse, NP), d.h. die stark innervierte Steuerdrüse für das gesamte endokrine System, wird durch den gepunkteten Streifen am unteren Rand des Bildes dargestellt. Die Nervenimpulse, die horizontal (von links nach rechts) laufen, wirken schließlich auf die motorische Oberfläche (MS), deren Veränderungen (Bewegungen) unmittelbar wiederum von der sensorischen Oberfläche (SS) wahrgenommen werden, wie dies durch die „äußere" Bahn in Richtung der Pfeile angedeutet wird. Die vertikal laufenden Impulse (von oben nach unten) stimulieren die Hirnanhangdrüse, deren Aktivität Steroide in die synaptischen Spalte entläßt, was durch die geschwungenen Linien in den Zwischenräumen ausgedrückt sein soll. Sie modifizieren dadurch den *modus operandi* aller synaptischen Verbindungen, und folglich den *modus operandi* des gesamten Systems. Besonders hervorzuheben ist die doppelte Schließung des Systems, das nun rekursiv nicht nur das verarbeitet, was es „sieht", sondern auch die Tätigkeit seiner eigenen Organe. Um diese zweifache Schließung noch deutlicher zu machen, schlage ich vor, die Zeichnung des Bildes 18 so um ihre beiden kreissymmetrischen Achsen zu wickeln, daß die künstlichen Grenzen verschwinden und ein Torus, wie in Bild 19, entsteht.

Hier wird der „synaptische Spalt" zwischen den motorischen und den sensorischen Oberflächen durch den gestreiften Meridian in der Mitte vorne, die Hirnanhangdrüse durch den gepunkteten Äquator abgebildet. Dies zeigt, so meine ich, *in nuce* die funktionale Organisation eines lebenden Organismus. Die Rechenprozesse innerhalb dieses Torus unterliegen einer nicht-trivialen Einschränkung, die durch das Postulat der Kognitiven Homöostase formuliert wird:

> „Das Nervensystem ist so organisiert (bzw. organisiert sich selbst so), daß es eine stabile Realität errechnet."

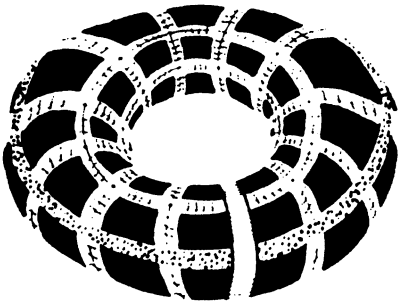


Bild 19

Dieses Postulat fordert „Autonomie", das heißt „Selbst-Regelung", für jeden lebenden Organismus. Da die semantische Struktur von Substantiven mit dem Präfix „selbst-" transparenter wird, wenn dieses Präfix durch das Substantiv ersetzt wird, wird der Ausdruck „Autonomie" synonym mit dem Ausdruck „Regelung der Regelung". Und genau dies leistet der doppelt geschlossene, rekursiv rechnende Torus: Er regelt seine eigene Regelung.

Bedeutung

Es mag in einer Zeit wie der unseren seltsam anmuten, Autonomie zu fordern, denn Autonomie bedeutet Verantwortung: Wenn ich selbst der einzige bin, der entscheidet, wie ich handle, dann bin ich für meine Handlungen verantwortlich. Da die Regel eines der populärsten Spiele, das man heute spielt, darin besteht, jemand anderen für *meine* Handlungen verantwortlich zu machen – der Name dieses Spiels lautet „Heteronomie" –, führen meine Überlegungen, soweit ich sehe, zu einer höchst unpopulären Auffassung. Ein Verfahren, diese Auffassung unter den Teppich zu kehren, besteht darin, sie bloß als einen erneuten Versuch der Rettung des „Solipsismus" zu betrachten, d.h. der Ansicht, daß diese Welt nur in meiner Vorstellung existiert, und daß die einzige Realität nur das sich etwas vorstellende „Ich" ist. In der Tat habe ich genau dieses vorhin festgestellt, ich habe dabei jedoch nur von einem einzigen Organismus gesprochen. Die Situation ist völlig anders, wenn es zwei davon gibt, wie ich mit Hilfe des in Bild 20 gezeichneten Herrn mit der Melone auf dem Kopf demonstrieren möchte.

Dieser Herr besteht darauf, daß er die einzige Realität ist, und daß alles übrige nur in seiner Vorstellung existiert. Auch dieser Herr kann jedoch nicht leugnen, daß das von ihm imaginierte Universum mit Erscheinungen bevölkert ist, die ihm selbst durchaus nicht unähnlich sind. Er muß folglich zugeben, daß auch diese Erscheinungen selbst darauf bestehen könnten, daß *sie* die einzige Realität sind, und daß alles übrige lediglich ein Produkt ihrer Einbildung ist. In diesem Falle aber ist das von ihnen imaginierte Universum mit Erscheinungen bevölkert, zu denen auch *er*, d.h. der Herr mit der Melone auf dem Kopf, gehören muß. Nach dem Prinzip der Relativität, das eine Hypothese ablehnt, die für zwei Phänomene zusammen nicht gilt, obwohl sie für jedes der beiden Phänomene allein zutrifft – Erdbewohner und Venusbewohner mögen beide darin übereinstimmen, daß sie behaupten, der Mittelpunkt des Universums zu sein, ihr Anspruch zerfällt aber, wenn sie aufeinandertreffen –, löst sich auch der solipsistische Standpunkt auf, sobald ich neben mir noch einen weiteren autonomen Organismus erfinde. Da das

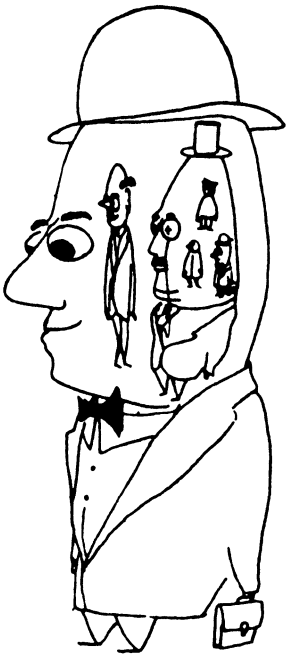


Bild 20

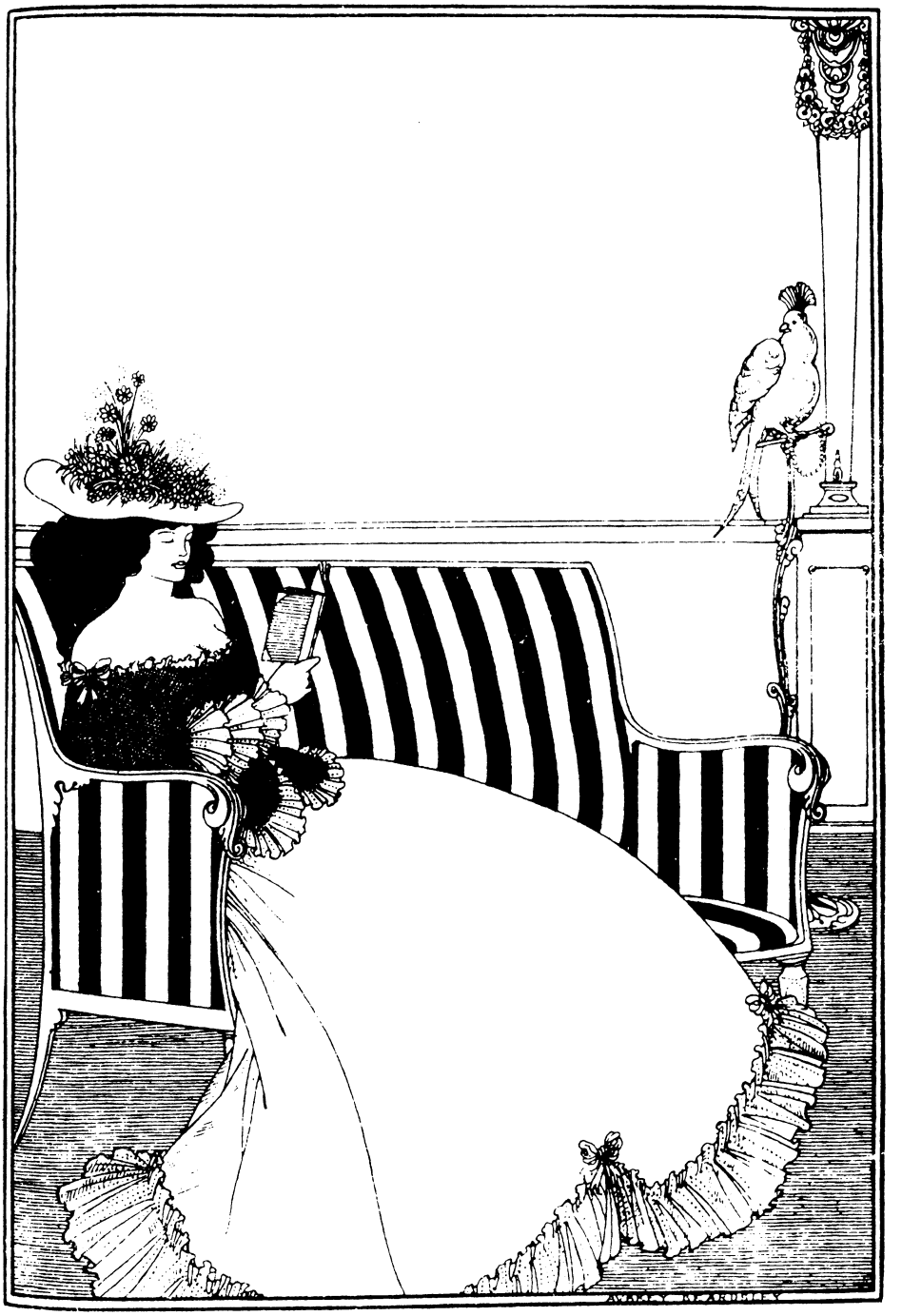
Prinzip der Relativität aber logisch nicht notwendig ist, noch auch eine Behauptung darstellt, die als wahr oder falsch zu erweisen ist, ist hier besonders hervorzuheben, daß der entscheidende Punkt, um den es geht, darin liegt, daß es mir freisteht, dieses Prinzip anzunehmen oder zu verwerfen. Wenn ich es ablehne, dann bin ich der Mittelpunkt des Universums, meine Wirklichkeit sind meine Träume und meine Alpträume, meine Sprache ist ein Monolog, meine Logik eine Monologik. Wenn ich das Prinzip akzeptiere, kann weder ich noch auch ein anderer den Mittelpunkt des Universums bilden. Es muß wie im heliozentrischen System etwas Drittes geben, das den zentralen Bezugspunkt bildet. Dies ist die Relation zwischen Du und Ich, und diese Relation heißt IDENTITÄT:

Realität = Gemeinschaft.

Worin liegen nun die Konsequenzen alles dieses für Ethik und Ästhetik?

Der ethische Imperativ: Handle stets so, daß die Anzahl der Wahlmöglichkeiten größer wird.

Der ästhetische Imperativ: Willst du sehen, so lerne zu handeln.


AUBREY BEARDSLEY

Bibliothekare und Technik: eine Mesalliance? *

[Eine Antwort]

Als mir Mrs. Howe vor einigen Tagen von ihren Beweggründen erzählte, mich zu dieser Veranstaltung einzuladen, war ich recht fasziniert von ihrer beharrlichen Weigerung, mir einen Titel für meinen Vortrag zu nennen. Sie bat mich vielmehr, auf eine Frage einzugehen, die die meisten von Ihnen persönlich sehr zu beschäftigen scheint, die Frage nämlich, welchen Einfluß die Technik auf die Zukunft Ihrer beruflichen Tätigkeit haben wird. Ich will mich gerne bemühen, Ihnen so gut wie möglich zu antworten, weil ich Bücher liebe und auch nie aufhören kann, Fragen zu stellen.

Lassen Sie mich Ihnen zunächst versichern, daß ich mir keinen günstigeren Augenblick dafür vorstellen kann, um ein Klima des Verständnisses zu erzeugen, in dem Ihre dringenden beruflichen Bedürfnisse in fruchtbarer Weise auf die Entwicklung der Technik Einfluß nehmen können. Ich meine nämlich, daß Sie auf keinen Fall abwarten sollten, was die Technik *mit Ihnen* tun wird – verzeihen Sie mir, die Techniker würden natürlich sagen *„für Sie"* –, sondern daß *Sie* vielmehr den Technikern sagen müssen, was diese für Sie tun sollen.

Mit anderen Worten, ich schlage vor, Mrs. Howes ursprüngliche Frage: „Die Technik: wie wird sie die Bibliothekare verändern?" umzukehren und nunmehr zu fragen: „Die Bibliothekare: wie werden sie die Technik verändern?"

Die Antwort auf diese Frage wird natürlich von Ihrer Kenntnis dessen abhängen, was technisch getan werden kann. Noch wichtiger aber: sie wird abhängen davon, wie Sie als Bibliothekare sich selbst als Teil ihrer Kultur in der Zukunft sehen, und welche Rolle Sie der Technik in diesem Ihren projizierten Selbstbild übertragen wollen.

Technik

Im Gegensatz zu der allgemeinen Überzeugung, daß wir heute in einem Zeitalter der Technik leben, behaupte ich, daß wir in einem Zeitalter der Technokratie leben. Damit meine ich, daß wir – hoffentlich nur für gewisse Zeit – unsere Verantwortung abgelegt haben, eine Technik zu fordern, die unsere gegenwärtigen Probleme löst. Stattdessen gestatten wir der bestehenden Technik, Probleme zu erzeugen, die sie lösen kann. So wird uns z. B. erzählt, daß die neuartigen Keramiktöpfe, die man aus dem Gefrierschrank auf den Herd holen kann, ein Nebenprodukt des Apollo-Raumprogramms sind. Ich würde lieber eine Welt sehen, in der das Apollo-Raumprogramm ein Nebenprodukt der Erzeugung solch nützlicher Pfannen und Töpfe wäre.

---

* Dieser Aufsatz geht zurück auf einen Vortrag vom 24. Juli 1970 im Rahmen des Library Institute, University Extension, The University of Wisconsin, Madison, Wisconsin.

In ähnlicher Weise kann ich mir eine Welt vorstellen, in der die für die Lösung der grundlegenden Probleme Ihres Berufs angemessene Technik erzeugt wird, und in der die Lösungen einiger Randprobleme der Bibliothekswissenschaft nicht nur ein Nebenprodukt industrieller Computerentwicklungsprogramme sind. Wie sehe ich diese grundlegenden Probleme?

Bibliothekare

Es sind zweierlei Funktionen, in denen ich den Bibliothekar in der Gesellschaft der Zukunft tätig sehe: einmals als Hüter von Büchern, zum anderen als Geburtshelfer für all jene, die neue Einsichten und Ideen in die Welt setzen wollen.

Diese zwei Möglichkeiten scheinen einer recht oberflächlichen Unterscheidung zu entspringen, solange man sich nicht klarmacht, daß jährlich etwa zwei Millionen Forschungsberichte aus allen Bereichen der Wissenschaft und Technik veröffentlicht werden, und daß sich diese Berichte auf etwa 30.000 verschiedene Fachzeitschriften verteilen. In diesem Jahrhundert werden sich diese Zahlen alle 10 bis 15 Jahre verdoppeln, ihre Zuwachsrate wird mehr als zweimal so hoch sein wie die der Erdbevölkerung (National Academy of Science 1969). Dies bedeutet, daß die Chancen eines Benutzers der Bibliothek der Zukunft, die Quelle für seine Erleuchtung konkret benennen zu können, sich mit erschreckender Geschwindigkeit vermindern. Er wird vielmehr seine Bitte um ein bestimmtes Buch oder Dokument zunächst zu einer Bitte um Titel geeigneter Bücher oder Dokumente umformulieren und fragen: „*Wo* findet sich die Antwort auf meine Frage?" Da ein Bibliotheksbenutzer aber primär daran interessiert ist, eine Antwort auf seine Frage zu erhalten, und nur sekundär daran, wo er diese finden kann, wird er am Ende immer fragen: „*Wie* lautet die Antwort auf meine Frage?"

Ich weiß, daß Sie sich der gegenwärtigen Verschiebung von der Dokumentenanforderung zur Titelanforderung sehr wohl bewußt sind, und daß Sie diesen Wünschen entsprechen, indem Sie umfangreiche Indexiersprachen, Kataloge mit Querverweisen, und komplizierte Verfahren der Herstellung von Kurzfassungen ("abstracts") entwickeln. Auf der anderen Seite bin ich auch überzeugt, daß Sie sich der Nachteile dieser Verfahren sowie ihrer Grenzen bewußt sind, die in erster Linie begrifflich-theoretischer und nicht technischer Natur sind. Vielleicht amüsiert es Sie zu erfahren, daß all diese höchst ausgefeilten Suchtechniken es nie ermöglichen würden zu entdecken, daß Max Planck die Quantenmechanik geschaffen hat: Der Ausdruck „Quantum" erscheint weder im Titel noch im Text seiner revolutionären Arbeiten (M. Planck 1900a; 1900b). Oder fühlt man sich nicht zu Recht entmutigt, wenn man eine Matrix der Querverweise zu erstellen versucht? Wenn man auch nur Querverweise k-ter Ordnung für D Dokumente möchte, hat die Matrix D^k Eintragungen, und dies bedeutet, daß man einem Katalog mit Querverweisen auch nur zweiter Ordnung pro Jahr nicht weniger als vier Billionen Eintragungen hinzufügen muß, wovon die meisten leer sind —aber als leer eingetragen werden müssen!

Obwohl unsere Techniker sich eifrig bemühen, Sie zur Investition gewaltiger Summen für teure Anlagen zu überreden, die als Heftpflaster die sachlichen Gebrechen nur verdecken, müssen Sie sich früher oder später auf die bereits erwähnte Wende einstellen, daß der Nutzer nämlich nicht mehr fragt „*Wo* ist die Antwort ...?", sondern „*Wie* lautet die Antwort ...?"

Mit anderen Worten: Ich behaupte, daß die Benutzer künftiger Bibliothekssysteme kein Interesse daran haben werden, irgendwelche Dokumente zu bekommen, die die Antwort auf ihre Frage enthalten oder auch nicht enthalten, sondern fordern werden, direkten Zugang zum *semantischen Inhalt* dieser Dokumente zu bekommen, und daß es ihnen gleichgültig sein wird, ob die Antwort des Systems ein wörtliches Zitat aus einem bestimmten Dokument ist oder eine dem Zitat äquivalente Paraphrase. Der Benutzer wird die Tatsachen verlangen und kein Interesse dafür zeigen, wie diese beschrieben sind, solange die Beschreibung korrekt ist und seinen Bedürfnissen entspricht.

Herausforderung

Das ist nun eine in der Tat sehr ernste Herausforderung, und sie kann nicht von den Bibliothekaren allein bewältigt werden. Ich füge sofort hinzu, daß dieser Herausforderung auch von keiner einzigen Wissenschaft alleine begegnet werden kann. Sie erfordert die Kooperation undogmatischer Fachleute in einem breiten Spektrum von Wissenschaften, um die Art von Systemen aufzubauen, die von Ihnen verlangt werden wird.

Ich hoffe, daß Sie erkennen, daß meine Behauptungen nicht die Idee der Bibliothek als eines Zentrums der Wissensgewinnung in Frage stellen. Was ich aber tatsächlich in Frage stelle, ist die Idee des Buches – bzw. der mit ihm verwandten Formen der Dokumentation – als des grundlegenden Vehikels der Wissensgewinnung.

Wenn ich mir mit diesen beiden Aussagen selbst zu widersprechen scheine – eine Bibliothek ist ja nichts anderes als eine *bonorum librorum copia*, ein „Schatzhaus guter Bücher" –, dann nur für den Fall, daß man auf dieser sehr engen Definition besteht. In jedem Falle zeigt diese Definition aber, wie stark unsere Kultur Bücher als Träger des gedruckten Wortes mit Behältern der Weisheit und des Wissens identifiziert – eine Überzeugung, die aus unserem jüdisch-christlichen Erbe stammt: „Und der Herr gab mir zwei Tafeln aus Stein, beschrieben mit dem Finger Gottes" (Deuteronomium 9/10). Beachten Sie, daß das Wort Gottes geschrieben ist und gelesen werden kann; außer für Moses wird es nicht gesprochen und kann daher nicht gehört werden.

Konfusion

Vielleicht ist die Ausdehnung unserer Verehrung der Heiligen Schrift, die das Wort Gottes *ist*, auf andere Schriften, die bloße Repräsentationen von Tatsachen oder Ideen sind, der Ursprung der Verwirrung, die ein Objekt mit dem Symbol identifiziert, das es repräsentiert. Sobald diese Gleichsetzung nicht mehr bloß als Konfusion angesehen werden kann, sondern zu einer Weise des Denkens geworden ist, handelt es sich um ein Symptom der Schizophrenie. Ein Patient, der gefragt wird „Wieviel ist 5 × 5?", kann eine detaillierte Beschreibung seiner Wohnung liefern, denn er wohnt zufällig im Haus Nr. 25.

Während diese Verwischung von Tatsachen und ihren Beschreibungen in ekstatischen, religiösen, patriotischen und pathologischen Zuständen akzeptiert werden kann, ist es empfehlenswert, sie nicht kritiklos zu adoptieren, denn dann wären wir nicht mehr in der Lage, eine Aussage (eine Beschreibung) als wahr

oder falsch zu beurteilen. Tatsachen sind, wie sie sind: sie sind weder wahr noch falsch. Nur die Beschreibungen von Tatsachen sind wahr oder falsch.

Nun lassen sich für diese Konfusion mildernde Umstände ins Treffen führen: jede Beschreibung ist selbst wieder eine Tatsache. Was aber für jedes Werkzeug gilt, das zwar ein Ding ist, aber einem außerhalb seiner selbst liegenden Zweck dient, gilt auch für eine Beschreibung, die zwar eine Tatsache darstellt, aber einen außer ihr liegenden Zweck erfüllt. Natürlich gibt es auch Menschen unter uns, die Bücher wegen der Schönheit ihres Einbands oder ihres Drucks sammeln, und natürlich gibt es die Bücher der Dichtung, deren Tatsachen eben er-dichtet, Fiktionen, sind.

Da der Benutzer unserer Bibliothek der Zukunft Tatsachen zu erfahren wünscht, müssen wir ihm „100 Grad Celsius" sagen, wenn er den Siedepunkt des Wassers wissen will, und ihm *Lady Chatterley's Lover* geben, wenn er um biographische Details aus dem Leben dieser bezaubernden Dame bittet.

Es stellt sich also die Frage nach Aufbau und Arbeitsweise eines Systems, in dem Sie die Doppelrolle des Hüters von Büchern sowie des Geburtshelfers für neue Ideen und Einsichten spielen können und so die Idee der Bibliothek als eines Orts bewahren, an dem Wissen gewonnen werden kann.

Dies aber ist gleichbedeutend mit zwei fundamentalen Fragen:

1. „Wie wird Wissen gewonnen?",
2. „Wie können wir diesen Prozeß mechanisieren?".

Kognition

Für die Beantwortung der ersten Frage müssen wir das Problem der Kognition lösen. Erst in den letzten Jahren aber haben wir überhaupt ein auch nur annäherndes Verständnis der Tiefe dieses Problems erreicht. Und gegenwärtig verfügen wir noch nicht über die Epistemologie, die Logik, die Mathematik und die Versuchsanordnungen, die eine solide Basis für das Verständnis dieser so rätselhaften Eigenschaft lebendiger Dinge bieten könnten. Aber schon das wenige, was wir in der letzten Zeit gelernt haben, hat viele unserer geheiligten Vorstellungen zutiefst verändert, und diese neuen Erkenntnisse erfordern wiederum eine radikale Abkehr von den bisherigen Ansichten von Systemen, die ihre Benutzer in die Lage versetzen sollen, das Wissen zu gewinnen, das sie benötigen.

Da ich gleich über die funktionale Organisation solcher Systeme sprechen werde, wie sie Ihnen hoffentlich in der Zukunft zur Verfügung stehen werden, möchte ich noch rasch den begrifflichen Rahmen skizzieren, innerhalb dessen wir nach Lösungen suchen müssen.

Das Problem der Kognition hat eine zweifache Wurzel: eine epistemologische und eine rechnerische. Da ein lebendiger Organismus eine autonome Ganzheit ist (Maturana 1970a), brauchen wir eine Epistemologie der „Autonomie". Mit Autonomie meinen wir, daß alle Entscheidungen über die Handlungen eines Organismus innerhalb seiner Haut getroffen werden. Ein lebendiger Organismus ist ein Universum in sich. Daraus folgt, daß ein vollständiger Formalismus für die Biologie – im Gegensatz zur Physik – in sich geschlossen sein muß. Das folgende Beispiel aus zwei komplementären Propositionen soll dies veranschaulichen:

1. Die Interpretationen der Wahrnehmungen eines Organismus determinieren seine Aktivität.
2. Die Aktivität eines Organismus determiniert die Interpretation seiner Wahrnehmungen.

Eine derartige zirkuläre Erklärung wird gewöhnlich als *circulus vitiosus* bezeichnet. Bei genauerem Hinsehen aber entdeckt man, daß es gerade diese Zirkularität ist, die das System in Gang hält. Denjenigen unter Ihnen, die mehr über die Zulässigkeit solcher geschlossener Formalismen erfahren möchten, empfehle ich die Arbeiten von Katz (1962), Löfgren (1968) und Brown (1972). Diese Autoren haben die logische Konsistenz und Vollständigkeit solcher Formalismen überzeugend nachgewiesen. Die exakte Abbildung dieser Kalküle auf identifizierbare funktionale Elemente lebendiger Organismen ist jedoch noch nicht gelungen, und zwar einfach deshalb, weil die Entwicklung der entscheidenden Experimente für diese Art von Ethologie über das Anfangsstadium noch nicht hinausgekommen ist. Auf der anderen Seite aber gibt es eine Fülle von Beobachtungen, die die Gültigkeit dieses Ansatzes bekräftigen. So berichten etwa Versuchspersonen, die über längere Zeit Brillen tragen, die das Sehfeld optisch umkehren, daß sie die Welt in den ersten Tagen auf den Kopf gestellt sehen, daß sich diese aber schrittweise wieder richtig einstellt, und zwar zuallererst in der direkten Körperumgebung im Abstand einer Armeslänge, dann im Bereich der Entfernung um einige Schritte, später noch weiter, bis schließlich, nach etwa zwei oder drei Monaten, das gesamte Sehfeld wieder so erlebt wird wie ohne die Umkehrbrillen. Dies deutet darauf hin, daß das Motorium das Sensorium organisiert. Dies kann, wie eine Versuchsperson berichtet, so weit gehen, daß der erste Schneefall des Jahres in einer ansonsten normal wahrgenommenen Umwelt nach oben fiel.

Angesichts solcher Beobachtungen sollten wir uns klar machen, daß wir nicht einfach von „Information" sprechen können, als wäre diese ein Gebrauchsgegenstand außerhalb des wahrnehmenden Bewußtseins. Die Welt enthält keine Information, die Welt ist, wie sie ist (von Foerster 1970a), Information über die Welt wird in einem Organismus durch seine Interaktionen mit der Welt erzeugt. Wenn nun aber einige der fortgeschrittenen Systeme für die Speicherung und Wiederbereitstellung von Dokumenten als Systeme für die Speicherung und Wiederbereitstellung von Information bezeichnet werden, gehen wir in eine gefährliche semantische Falle. Solche Systeme speichern Bücher, Magnetbänder, Mikrofiches oder andere Arten von Dokumenten, denn sie können natürlich keine „Information" speichern. Und eben diese Bücher, Bänder, Mikrofiches oder andere Dokumente werden wieder bereitgestellt, und sie liefern nur dann die gewünschte Information, wenn sie von menschlichen Augen betrachtet werden. Wenn man so die *Träger* potentieller Information mit der *Information* selbst vermengt, verschiebt man das Problem der Kognition bequem in den blinden Fleck des geistigen Sehfeldes.

Lassen Sie mich nun zu den rechnerischen Problemen der Kognition übergehen. Wir begegnen diesen schon auf der grundlegendsten Ebene, etwa wenn wir fragen: „Was konstituieren die sogenannten ‚Sinnesmodalitäten'?", und sie erstrecken sich bis zur Ebene der höheren mentalen Funktionen, wenn wir zum Beispiel fragen: „Was ist Gedächtnis?", „Was ist Lernen?" usw.

Vielleicht überrascht es Sie zu hören, daß wir die neuronalen Rechenprozesse noch nicht erklären können, die zu den Erfahrungen von Klang, Licht und Farbe,

Geruch und Geschmack usw. führen. Ursprünglich meinte man, daß stimulusspezifische Rezeptorzellen allein für diese unterschiedlichen Erfahrungen verantwortlich wären, wie z. B. die Rezeptoren für bestimmte Wellenlängen des elektromagnetischen Spektrums (die „Zapfen" der Retina), für bestimmte Lichtstärken (die „Stäbchen") oder für bestimmte molekulare Konfigurationen (die Geschmackspapillen und die Geruchsorgane) usw. Dies trifft aber nicht zu, denn keiner dieser Rezeptoren enkodiert in seiner Aktivität die physikalischen Ursachen dieser Aktivität. Die einzige Botschaft, die die Rezeptoren vermitteln, ist: ‚Es gibt so und so viel (aber nicht ‚was') an dieser Stelle meines Körpers." Da unsere Sinnesrezeptoren folglich nicht fähig sind, die Unterschiedlichkeit der physikalischen Agentien, die ihre Aktivität auslösten, zu übermitteln, ist die „prachtvolle Vielfalt ' unserer Erfahrungswelt, das ‚Was', ein Ergebnis der Verrechnung der von den Rezeptoren gelieferten Signale.

Semantische Rechenprozesse

Lassen Sie mich nun die Auffassung, daß Rechenprozesse und nicht Signalübertragung und Signalspeicherung den Kern der kognitiven Prozesse bilden, durch eine kurze Betrachtung des „Gedächtnisses" verdeutlichen. Denn genau das Mißverständnis dieser höheren mentalen Funktion als eines „Systems der Datenspeicherung" blockiert unsere Einsicht in die Art von Systemen, die wir für die Bewältigung der Herausforderungen der Zukunft brauchen.

Wäre unser Gedächtnis ein Datenspeicherungssystem, dann läßt sich leicht zeigen, daß wir für die Erklärung dessen, was wir wissen, ein Gehirn von der Größe einer dicht mit Nervenzellen vollgepackten Kugel mit einem Durchmesser von einer Meile haben müßten (von Foerster/Inselberg/Weston 1968). Hätte unser Gehirn aber diese Größe, dann würde es für den Prozeß der Wahrnehmung eines Löwen in seinem Sehfeld etwa zehn Jahre benötigen. Das könnte zwar dem Löwen dienlich sein, kaum aber dem Träger des Gehirns.

Um dies völlig klar zu machen, wollen wir für einen Augenblick annehmen, wir möchten die Multiplikation von Zahlen dadurch fehlerfrei machen, daß wir alle Produkte zweier Zahlen von bis zu n Ziffern in einer gedruckten Tabelle speichern. Die Länge L des Regals für die Aufbewahrung dieser Tabelle, gedruckt auf normalem Papier vom Format 21 X 27 cm, läßt sich leicht berechnen:

$$L = n \cdot 10^{(2n-6)} \text{ cm.}$$

Das ergibt für Produkte von Faktoren von 1 – 1000 (n = 3) ein Buch von 3 cm Dicke bzw. ein entsprechend bescheidenes Regal. Nun mag das zwar für ein Schulkind ausreichen, um aber heute mit normalen Geschäftsanforderungen fertigzuwerden, müssen wir schon zu zehnstelligen Zahlen greifen (n = 10). Dann wird die Tabelle 10^{15} cm lang, d. h. hundertmal die Entfernung Erde-Sonne oder etwa einen Lichttag. Ein Bibliothekar, der sich mit Lichtgeschwindigkeit bewegt, braucht im Schnitt einen halben Tag, um eine einzige Eintragung in dieser Tabelle nachzuschlagen.

Vergleichen Sie nun die Größe dieses Speichers mit einer Rechenmaschine, die in eine Hand paßt und genau das gleiche leistet (Abb. 1). Es genügen im Durchschnitt 30 Umdrehungen der Kurbel auf der Oberseite, und das gewünschte Ergebnis der Multiplikation zweier achtstelliger Zahlen erscheint in den Fenstern

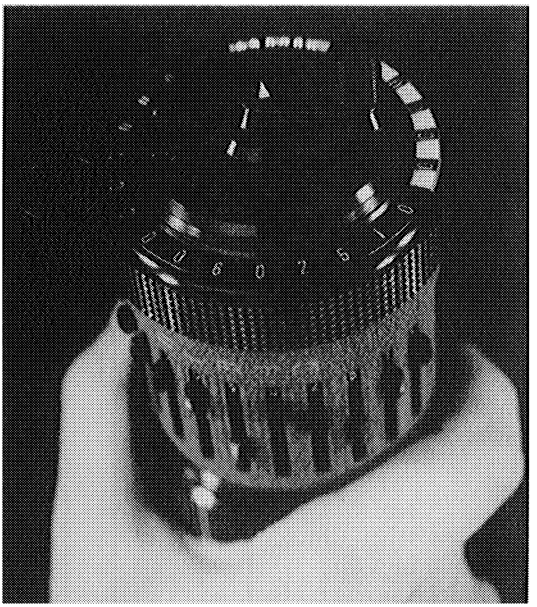


Abb. 1
„Curta", ein digitaler Handrechner, der Produkte bis zu 10^{16}-1 bewältigt

des Produktregisters. Dieses Gerät speichert klarerweise keine Daten, es rechnet *mit* Daten, in unserem Falle mit den Faktoren eines Produkts. Will man in diesem Fall überhaupt von „Speicherung" sprechen, dann kann sich dies nur auf die innere mechanische Struktur dieses Geräts beziehen, die sozusagen das Prinzip numerischen Rechnens „verkörpert".

Auch wenn wir nicht im einzelnen verstehen, wie das Nervensystem solche Rechenprozesse bewerkstelligt, erlaubt uns das Verständnis der dafür geltenden logischen Beziehungen, diese in der Struktur mechanischer oder elektronischer Systeme abzubilden, die nur die Operationen ausführen können, die ihre Struktur ihnen vorschreibt.

Die für Rechenprozesse im logisch-mathematischen Bereich notwendigen Beziehungen verstehen wir heute sehr gut, daher auch der Erfolg der Geräte, die diese Beziehungen verkörpern: der digitalen Großcomputer. Die Struktur semantischer Beziehungen aber, wie sie in der funktionalen und anatomischen Organisation unseres Gehirns verkörpert ist und uns auf andere reagieren und mit anderen durch Sprache und Verhalten interagieren läßt, wird erst jetzt erforscht und langsam verständlich. Bis vor kurzem waren auch die Linguisten nicht besonders hilfreich für die Lösung dieses Problems. Sie haben sich fast ausschließlich mit *Syntax* beschäftigt, d. h. mit den Regeln, nach denen Symbole zu wohlgeformten Ketten verknüpft werden können, *Semantik* dagegen, d. h. die Regeln, die diesen Symbolketten Bedeutung verleihen, war lange Zeit ein schmutziges Wort. Das ist aber vorbei, seitdem klargeworden ist, daß syntaktische Ambiguitäten im semantischen Bereich desambiguiert werden. Aufgrund des ständigen Fortschritts in diesem neuen Forschungsbereich, der „Psycholinguistik", wird es nunmehr möglich, den Begriff des Rechnens um das Rechnen im semantischen Bereich zu erweitern.

Da dieser Begriff des Rechnens im semantischen Bereich von ganz entscheidender Bedeutung ist, wenn es um die Entwicklung der Computerarchitektur für die Zentren der Wissensgewinnung der Zukunft geht, möchte ich Ihnen ein Beispiel für semantisches Rechnen geben, das wir P. Weston verdanken (Weston 1970a).

Weston hat die implizite Relationenstruktur jener Denkaufgaben untersucht, in denen am Anfang eine Geschichte in Form einer Menge scheinbar unverbundener Aussagen steht und daraufhin Fragen nach Einzelheiten gestellt werden, deren Beantwortung unmöglich scheint. Rätselfans bezeichnen sie als den „Smith-Robinson-Jones"-Typ.

Hier ist ein von Weston analysiertes Beispiel:

Ein Zug wird von drei Männern geführt, von Smith (S_1), Robinson (R_1) und Jones (J_1). Sie sind Lokomotivführer (E), Heizer (F), und Bremser (B), aber nicht notwendig in dieser Reihenfolge. Im Zug befinden sich drei Geschäftsleute mit den gleichen Namen, Mr. Smith (S_2), Mr. Robinson (R_2) und Mr. Jones (J_2). Für alle gelten die folgenden Tatsachen:

1. Mr. Robinson lebt in Detroit.
2. Der Bremser lebt genau in der Mitte zwischen Chicago und Detroit.
3. Mr. Jones verdient genau 20.000 Dollar im Jahr.
4. Smith hat den Heizer im Billiard geschlagen.
5. Der nächste Nachbar des Bremers, einer der Passagiere, verdient dreimal so viel wie der Bremser, der 10.000 Dollar im Jahr verdient.
6. Der Passagier, der den gleichen Namen hat wie der Bremser, lebt in Chicago.

Das alles ist gegeben. Daraufhin werden z. B. die folgenden Fragen gestellt:

„Wer ist der Lokführer?"

Oder:

„Welche Beziehung besteht zwischen Jones und dem Passagier, der den gleichen Namen hat wie der Lokführer?"

Usw.

Das sind nun alles anscheinend ganz unmögliche Fragen, ich möchte Sie aber doch gerne darauf hinweisen, daß genau diese Art von Fragen in Zukunft an Sie gestellt werden wird, und daß es Ihre Aufgabe sein wird, aus einer „Datenbasis", die der des Rätsels recht ähnlich ist, die Antworten herauszuholen.

Wie lassen sich nun diese Fragen beantworten? Weston hat sich gefragt, ob die in der Problembeschreibung dargestellte Situation in eine einzige komplexe Relationenstruktur übersetzt werden kann (oder nicht), die als Basis für alle von den Fragen geforderten Operationen dienen könnte. Er stellte fest, daß im gegebenen Fall die Relationenstruktur der Problembeschreibung auf fünf binären Relationen zwischen Elementen aus fünf verschiedenen Mengen beruht: der Menge der Zugleute, TM, mit den Elementen Smith, Robinson, Jones; der Menge der Tätigkeiten, J, = Lokführer, Heizer, Bremser; der Menge der Passagiere, P, der Orte, L der Gehälter, S, – mit ihren jeweils leicht bestimmbaren Elementen. Die fünf binären Relationen heißen „Gleichnamigkeit", n; „Tätigkeit", o; „schlägt", w; „wohnt in", r; und „verdient Geld", m. Diese Relationen heißen „binär", denn sie behaupten eine Beziehung zwischen zwei „Variablen", z. B. n (x, y), oder in Worten: „x hat den gleichen Namen wie y". Ebenso bedeutet m (x, y) „x verdient im Jahr y Dollar", usw.

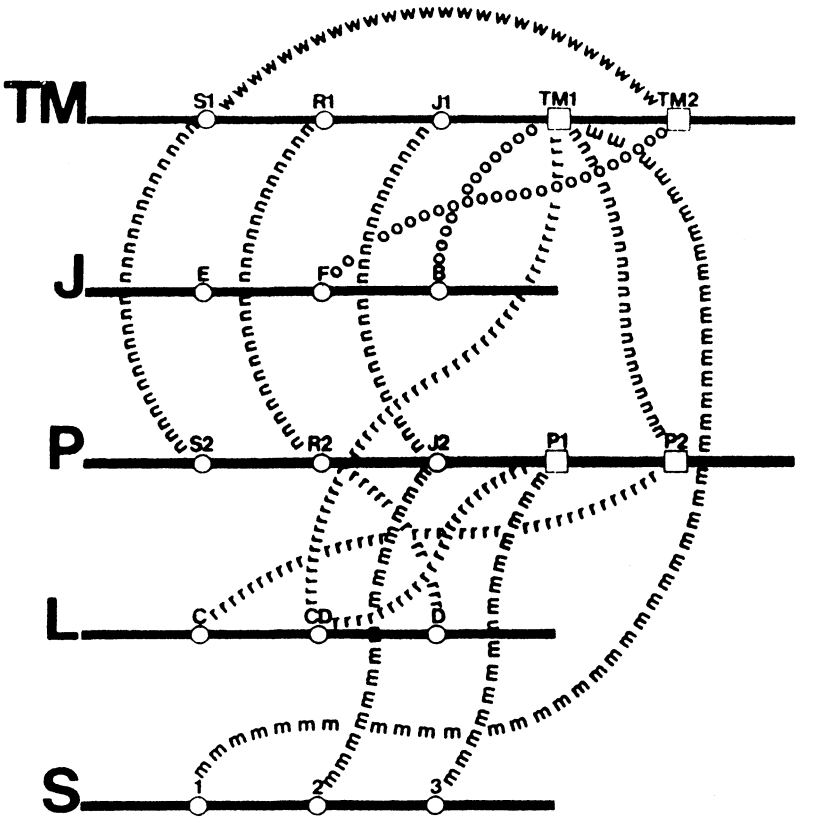


Abb. 2 Die semantische Struktur der „Smith-Robinson-Jones"-Problembeschreibung

Nach dieser Klärung läßt sich die gesamte Problembeschreibung tatsächlich durch eine einzige Relationenstruktur abbilden (Abb. 2). Mengen werden durch entsprechend bezeichnete waagerechte Linien dargestellt, Elemente als Punkte auf den zugehörigen „Mengenlinien", und Relationen durch die Buchstabenketten aus den Namen der Relationen, die die jeweiligen Elemente miteinander verbinden. Sie können ohne große Mühe die Problembeschreibung aus dieser Abbildung „ablesen" und umgekehrt.

Die Abbildung zeigt die „Datenbasis", auf der alle weiteren Berechnungen vorgenommen werden können. Beachten Sie bitte als erstes, daß diese Datenbasis der anderen Datenbasis äquivalent ist, die in der Problembeschreibung, also dem „Dokument" besteht, allerdings mit dem grundlegenden Unterschied, daß Sie nun das semantische Modell der Denkaufgabe direkt betrachten können, während es im Dokument undurchschaubar bleibt. Darüber hinaus läßt sich für diese Art von Darstellung ein „Algorithmus", d. h. eine Rechenregel, formulieren, mit der alle erforderlichen Rechnungen ausgeführt werden können. Wir brauchen uns um diese Rechnungen nicht weiter zu kümmern, wir können sie Maschinen überlassen.

Ich hoffe, daß Ihnen dieses Beispiel zumindest eine vage Vorstellung von dem vermittelt, was mit „Rechnen im semantischen Bereich" gemeint ist und was dies für den Benutzer der Zukunft bedeutet, der damit direkten Zugang zur semantischen Struktur der Datenbasis hat. Da das Bild eines „Dokuments" in der verteilten „Weisheit" dieser relationalen Datenbasis aufgegangen ist, kann der Benutzer an jedem gewünschten Punkt einsteigen, und seine Frage wird jeweils so „paraphrasiert", daß eine Antwort gegeben werden kann. Ist er zufriedengestellt, kann er wieder aussteigen, ist er es nicht, kann er weiter fragen. „*Die* Antwort" ist ohnehin eine Illusion. Antworten auf die Frage etwa „Wann wurde Napoleon geboren?" können lauten: „52 Jahre, bevor er auf St. Helena starb", „Eintausendsiebenhundertundneunundsechzig Jahre nach Christi Geburt", „Sieben Jahre vor der amerikanischen Unabhängigkeitserklärung" usw. Welche Art von Antwort Sie bevorzugen, hängt davon ab, was für ein Mensch Sie sind.

Es stellen sich nun zwei Fragen. Die erste ist: „Gibt es Maschinen, die die Dokumente in solche Datenbasen übersetzen können?", und die zweite: „Können Maschinen derartige Datenbasen verkörpern?". Auf die erste Frage läßt sich mit einem zurückhaltenden „Ja", auf die zweite dagegen mit einem eindeutigen „Ja" antworten. Weston (1970b) hat eine Programmstruktur mit Namen CYLINDER entwickelt, die die Abbildung eingebetteter Relationenstrukturen von beliebiger Tiefe erlaubt, und die Hardware für die Verwirklichung solcher Strukturen steht zur Verfügung.

Das zurückhaltende „Ja" auf die erste Frage ist jedoch nicht durch einen Mangel an Zuversicht, an Wissen oder an technischer Kapazität begründet, in der gegenwärtigen Phase ist es mehr als alles andere der Mangel an Geld, der uns aufhält. Das scheint nun ziemlich banal zu sein: wir alle haben zu wenig Geld! Daß dieses Problem in unserem Fall aber keineswegs so banal ist, wird gleich deutlich werden, wenn ich Ihnen eine kleine Skizze der verschiedenen Ansätze zur Mechanisierung von Systemen gebe, mit denen ein Benutzer einen lebendigen und informativen Dialog führen kann.

Computer für die Semantik

Es gibt zwei bedeutsame Verfahrensansätze für die Konstruktion von Computersystemen, die auf Fragen antworten können, die der Benutzer in seiner Umgangssprache gestellt hat. Ein Ansatz trägt den angemessenen Namen „Frage-Antwort-System", kurz QA, er ist historisch der Vorläufer des zweiten Typs, den ich mit „Kognitives Gedächtnis", kurz CM, bezeichne. Diese zwei Systeme sind zwar ähnlich in dem Sinne, daß sie beide Formulierungen in einer Sprache akzeptieren bzw. anbieten, die ein Benutzer seinerseits akzeptieren bzw. anbieten würde, sie unterscheiden sich aber darin, daß die Datenbasis des QA-Systems ein unveränderbarer „Kodex" ist, der nur vom Systemprogrammierer verändert werden kann, wenn neue oder andere Daten verfügbar sind, daß sich die CM-Datenstruktur jedoch nach jeder Interaktion verändert und so immer wieder das Ergebnis von Interaktionen in sich aufnimmt, um seine strukturelle Komplexität zu vergrößern. Dieser Unterschied kann prägnant dadurch benannt werden, daß man QA-Systeme als „maschineninvariant", CM-Systeme als „benutzeradaptiv" bezeichnet.

Im folgenden möchte ich versuchen, die „Anatomie und Physiologie" dieser beiden Arten von Computersystemen zu skizzieren, allerdings mit der ausdrücklichen Bitte um Nachsicht dafür, daß ich in der mir verbleibenden Zeit der Kom-

plexität, Subtilität und dem Erfingungsreichtum keineswegs gerecht werden kann, die für die Arbeitsweise solcher Systeme charakteristisch sind.

QA-Systeme

Im Prinzip ist hier die grundlegende Idee, die Struktur des Dokuments, wie es vorliegt, im großen und ganzen beizubehalten, indem man die N primitiven Sätze, die die Datenbasis bilden, als die Axiome A_1, A_2, A_3, ... A_n eines logischen Systems betrachtet und jede eingegebene Frage als ein Theorem, das als wahr oder falsch erwiesen werden kann. Die Frage z. B. „Darf ich X tun?" wird behandelt als „Beweise die Gültigkeit von X". Ist X gültig, dann ist die Antwort „Ja", sonst ist sie „Nein". Da eine so simple Antwort für den Benutzer aber sicherlich unbefriedigend ist – er möchte ja zumindest wissen, warum –, geben fortgeschrittenere QA-Systeme auch die „Gründe" für ihre Antwort an, indem sie auf Axiome in ihrer Datenbasis verweisen, die für die Ableitung der Antwort benutzt worden sind.

Der Computerausdruck eines typischen Dialogs mit einem QA-System, das auf die Straßenverkehrsregeln des Staates Illinois spezialisiert ist, findet sich in Abbldung 3.

```
F. DARF MAN EIN AUTO AUF DER RECHTEN SEITE ÜBERHOLEN?
A. NUR AUF EINER STRASSE ODER AUTOBAHN MIT ZWEI ODER MEHR IN FAHRT-
   RICHTUNG UNGESPERRTEN FAHRSTREIFEN.
F. WIE ALT MUSS MAN IN ILLINOIS SEIN, UM EINEN FÜHRERSCHEIN ERWERBEI
   ZU KÖNNEN?
A. DAS MINDESTALTER FÜR DEN FÜHRERSCHEIN IST 18 JAHRE. AUSGENOMMEN
   SIND PERSONEN, DIE 16 und 17 JAHRE ALT SIND UND EINEN ANERKANNTEI
   FAHRSCHULKURS ERFOLGREICH ABGESCHLOSSEN HABEN.
```

Abb. 3 Computerausdruck eines Dialogs mit dem R-2 System

Abbildung 4 zeigt übersichtsweise den inneren Aufbau dieses QA-Systems – es heißt „R 2" = "Rules of the Road" <Straßenverkehrsregeln>, ebenso wie das Illinois Driver Manual <Illinois Kraftfahrer-Handbuch>, dessen Vorschriften die Axiome von R 2 bilden (Bielby et al. 1970). Der Benutzer tippt Sätze und Fragen in englischer Sprache über ein Computerterminal in das System ein. Diese Sätze werden sodann vom „Parser" syntaktisch analysiert, und das Analyseergebnis wird an den "Semantic Interpreter" zur Desambiguierung weitergegeben, der zusammen mit der in der Einheit "Memory" gespeicherten Datenbasis die notwendigen semantischen Analysen und Umwandlungen ausführt, die einen Vergleich der Eingabe mit den Eintragungen der Datenbasis ermöglichen. Daraufhin werden die für den Beweis der gestellten Fragen erforderlichen Rechenprozesse eingeleitet ("Deducer"), die Ergebnisse in die Sprache des Benutzers übersetzt ("Transformer") und schließlich auf dem Terminal des Benutzers ausgedruckt.

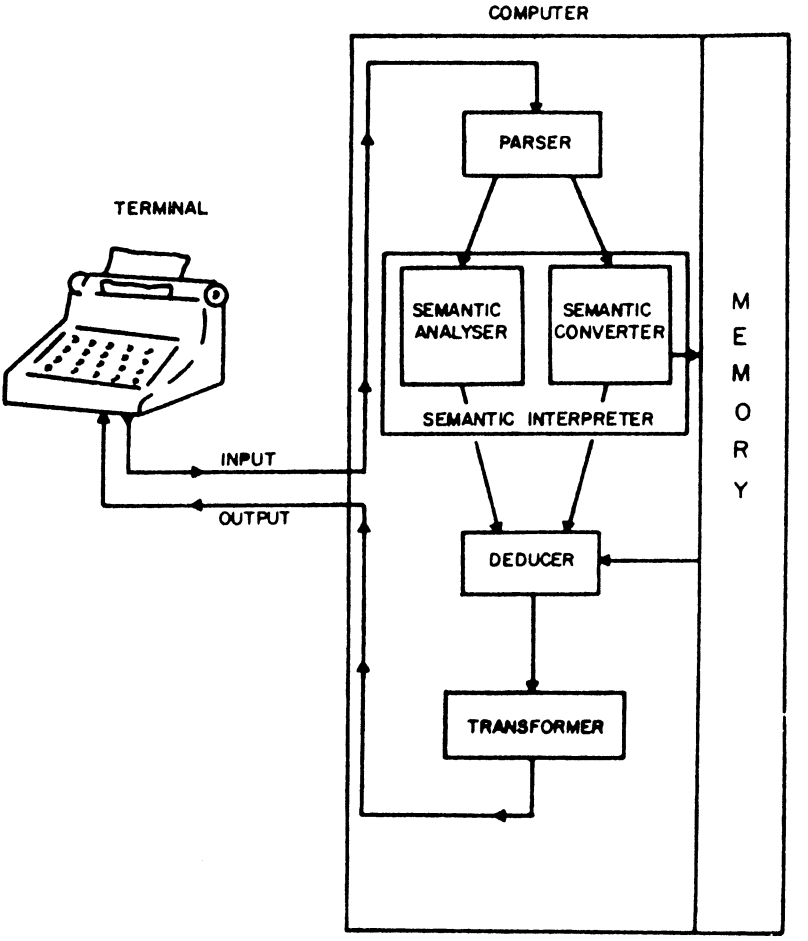


Abb. 4 Schematische Darstellung des R-2 Frage-Antwort-Systems

CM-Systeme

In QA-Systemen besteht der Großteil der semantischen Rechenprozesse in der Analyse jeder einzelnen Frage und dem Vergleich des Analyseergebnisses mit einer bereits vorhandenen Aufzeichnung des Originaldokuments, also in unserem Fall mit der Beschreibung der SRJ-Denkaufgabe (s. o.). CM-Systeme dagegen erledigen die semantischen Berechnungen bereits, wenn die Datenbasis aufgebaut wird, wenn also z. B. die spezifische relationale Struktur der SRJ-Denkaufgabe, wie sie Abb. 2 zeigt, konstruiert wird. Auch wenn die hier schon am Anfang gegebene Komplexität die von einer eingegebenen Frage geforderten Rechenprozesse nicht wesentlich verringert, erlaubt dieses System aber bei jeder Interaktion die Veränderung seiner Datenbasis, indem es Verknüpfungen beseitigt, die *implizit*

gemacht worden sind, und zwar durch die aus der schon existierenden Datenstruktur berechneten Deduktionen und Schlußfolgerungen. Es handelt sich hier um *Explikationen* im eigentlichen Wortsinn, und es ist zu beachten, daß das bessere Verständnis des konkreten Falles, das dem Benutzer durch die Antwort des Systems vermittelt wird, im System dadurch gespiegelt wird, daß dieses nunmehr besser organisiert ist, um künftige Fragen der gleichen Art zu bewältigen. Da solche Prozesse Modellen der Kognition am nächsten kommen (von Foerster 1970a, 1969), scheint der Name „Kognitives Gedächtnis" diese Systeme treffend zu benennen. Da diese Prozesse außerdem spezifische Rechenvorgänge erfordern, deren Programme selbst Teil der Datenbasis sein können, werden Organisationsformen, bei denen die Unterscheidung zwischen Datenspeicher und Programmen verwischt wird, „Kognitive Datenbasen", kurz CDB, genannt.

Abbildung 5 stellt den inneren Aufbau eines CM dar, und Tabelle 1 erklärt die verwendeten Symbole. Seine Arbeitsgänge lassen sich in folgenden Schritten veranschaulichen:

(0) Über eine Konsole wird eine Symbolkette (die Frage) (SS_{IN}) in das System eingegeben. Daraufhin wird aus der Menge A der Interface-Operatoren eine geeignete Teilmenge

$$\alpha \underset{IN}{\equiv} \{a_1\}$$

aktiviert, um die gegebene Kette in die Datenstruktur zu übersetzen. Dies hat zwei Konsequenzen:

(1) (i) Gewisse Bereiche des Datenzentrums werden modifiziert $[(D, E)_o \rightarrow (D, E)_1]$, und eine geeignete Teilmenge primärer Operatoren

$$\pi_1 = \{P_i\}_1$$

wird aktiviert.
(ii) Diese bilden nun in Verbindung mit der geänderten Datenstruktur aus den Elementarprogrammen E_2 eine Menge neuer sekundärer Operatoren

$$\omega_1 = \{O_1\}_1,$$

die mit einer Teilmenge von Interface-Operatoren

$$\overline{A} = \{\overline{a}_i\}$$

übereinstimmen oder nicht.

(2) Ist dies nicht der Fall, ergeben sich zwei Konsequenzen:
(i) Gewisse Bereiche des Datenzentrums werden modifiziert $[D, E)_1 \rightarrow (D, E)_2]$, und eine geeignete neue Teilmenge primärer Operatoren

$$\pi_2 = \{p_i\}_2$$

wird aktiviert.
(ii) Diese bilden in Verbindung mit der geänderten Datenstruktur aus den Elementarprogrammen E_1 eine Menge sekundärer Operatoren

$$\omega_2 = \{O_i\}_2,$$

die mit einer Teilmenge von Interface-Operatoren

$$\overline{A} = \{\overline{a}_i\}$$

übereinstimmt oder nicht.

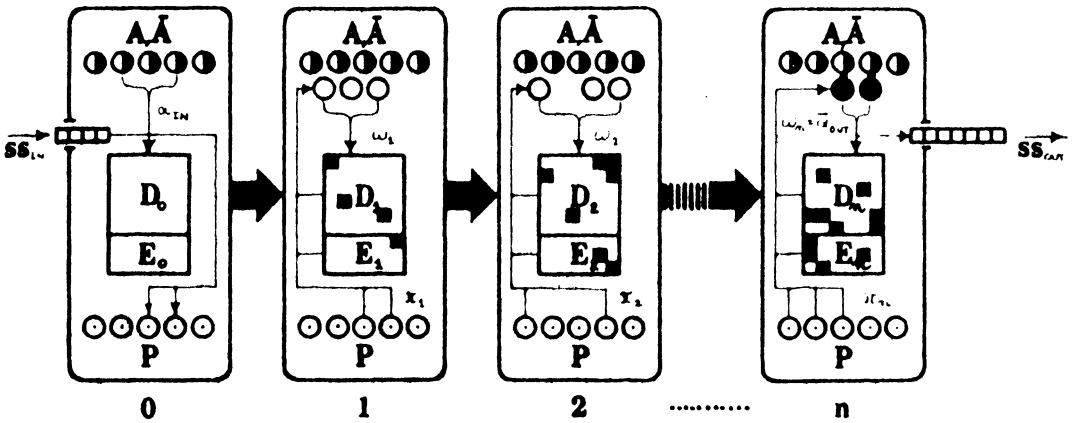


Abb. 5 Symbolische Abbildung aufeinanderfolgender Rechenschritte in einem CM-System

Tabelle I Erklärung der in Abb. 5 sowie in der im Text gegebenen Beschreibung eines CM-Systems verwendeten Symbole

Graphisches Symbol	Name	Symbol	Index-Bereich	Menge	Teil-Menge	Relationen-matrix
◖	Interface-Operatoren SS → CDB	a_i	$i = I \to n_a$	$A = \{a_i\}$	$\alpha = \{a_i\}$	$M(\alpha)$
◗	Interface-Operatoren CDB → SS	$\bar{a}_i = (a_I^{-1})$	$i = I \to n_a$	$\bar{A} = \{\bar{a}_i\}$	$\bar{\alpha} = \{\bar{a}_i\}$	$M(\bar{\alpha})$
⊙	Primäre Operatoren	P_i	$i = I \to np$	$P = \{P_i\}$	$\pi = \{P_i\}$	$M(\pi)$
○ ●	Sekundäre Operatoren	O_i	$i = I \to n_0$	$O = \{o_i\}$	$\omega = \{o_i\}$	$M(\omega)$

D ..... Datenstruktur
E ..... Elementprogramme
} Kognitive Datenbasis, CDB

τ ..... Zahl der Zeitintervalle $(\Delta t)_0 : f = \tau\,(\Delta t)_0$, $\tau = 0, 1, 2$

$M(\epsilon)_\tau = |m(\epsilon)_{i,j}|_\tau$ ... Matrix der Relationenstruktur $M(e_i, e_j)$, die für $a_\#$ Operatoren einer Teilmenge $\epsilon = \{e_i\}$ gilt, die während des Zeitintervalls τ aktiv sind.

(3) Ist dies nicht der Fall, dann wird der Prozeß rekursiv wiederholt mit

$(D, E)_\tau \to (D, E)_{\tau+1}$
$\pi_\tau > \pi_{\tau+1}$
$\omega_\tau > \omega_{\tau+1}$

bis eine Teilmenge von sekundären Operatoren ω_n errechnet ist, die mit einer Teilmenge von Interface-Operatoren

$\omega_n = \{O_i\} = \{\bar{a}_1\} = \bar{\alpha}$

übereinstimmt.

Diese Operatoren $\{a_i\}$ übersetzen nun den passenden Bereich der Datenstruktur in eine Symbolkette (SS_{OUT}), die auf der Konsole ausgedruckt wird. Im Falle der „Smith-Robinson-Jones"-Denkaufgabe etwa antwortet das System auf die zwei Fragen (s. o.) „Smith" bzw. „Sie leben am gleichen Ort".

Ökonomische Aspekte

Tabelle II enthält nach verschiedenen Größeneinheiten angeordnet die Menge des Materials, das bewältigt werden muß. In den Zeilen findet sich die Anzahl der Bausteine, aus denen die einzelnen kleineren bzw. größeren Einheiten zusammengesetzt sind. Ein Zeitschriftenartikel etwa besteht (im Schnitt) aus 10 Seiten, enthält 10.000 Wörter, umfaßt einen Block von 500.000 bits, usw., die Zahlen in den fettgedruckten Quadraten drücken die primären numerischen Relationen aus, von denen alle übrigen abgeleitet sind (jeweils abgerundet). Sie sind so gewählt, daß sie mit anderen Schätzungen vergleichbar sind (National Academy of Science 1969).

Auf der Grundlage dieser Zahlen kann man nun zwei Fragen stellen:

„Wie hoch sind die Installationskosten?", und
„Wie hoch sind die Betriebskosten?"

von – vorhandenen oder geplanten – Computersystemen, die in der Lage sind, die Interaktionen mit ihren Benutzern für Datenbasen zu bewältigen, ob diese nun sehr kleine Dokumentensammlungen darstellen oder riesige Bibliotheken von der Größe der Library of Congress heute oder in 10 oder 30 Jahren.

Gewöhnlich richtet man sein Augenmerk bei der Beantwortung solcher Fragen auf die Kosten der Maschinen sowie ihrer Wartung und Instandhaltung. Dies ist die typisch „maschinenorientierte" Einstellung, die völlig von der Arbeit absieht, die in der Erzeugung und Nutzung der Dokumente liegt. Wie wir aber gleich sehen werden, enthüllt erst die „gesellschaftsorientierte" Einstellung die verborgenen Kosten der Nutzung eines Systems.

Ich werde nun aber zunächst auf konventionelle Weise vorgehen, um die Installationskosten zu ermitteln. Um ein Vergleichsmaß zu gewinnen, möchte ich drei Systeme erörtern: die bekannte Methode der Indexiersprache (IL), sowie die zwei vorhin besprochenen Systeme, das Frage-Antwort-System (QA) und das Kognitive Gedächtnis (CM).

Der einfachste Weg zu einer Schätzung der Installationskosten für ein Gesamtsystem besteht darin, sich die Kostenzuwachsrate (Δy in Dollar) anzuschauen, die sich ergibt, wenn seine Kapazität mit einer bestimmten Zuwachsrate von „Einheiten" (Δx) wächst. Als geeignetste Einheit dafür bietet sich ein einzelnes Heft einer Zeitschrift oder dessen Äquivalent, ein Buchkapitel, an.

Fügt man IL-Systemen neues Material hinzu, dann muß man neue Index-Wörter aus bereits indexierten Dokumenten herausziehen, QA- und CM-Systeme dagegen erfordern lediglich zusätzliche Bestandteile, um die erweiterte Datenbasis zu handhaben. Es ergibt sich also:

(i) für IL-Systeme

$$\Delta y = C_1 N \Delta x.$$

Tabelle II Konversion verschiedener Einheiten der Dokumentation

	Bits	Buch-staben	Wörter	Sätze	Seiten	Aufsätze	Hefte (Kapitel)	Bücher	kleine Bibliotheken	große Bibliotheken
1 Buchstabe	8	1								
1 Wort	48	6	1							
1 Satz	$6 \cdot 10^2$	75	12.5	1						
1 Seite	$5 \cdot 10^4$	$6 \cdot 10^3$	10^3	80	1					
1 Aufsatz	$5 \cdot 10^5$	$6 \cdot 10^4$	10^4	800	10	1				
1 Heft (Kapitel)	$5 \cdot 10^6$	$6 \cdot 10^5$	10^5	$8 \cdot 10^3$	100	10	1			
1 Buch	$2 \cdot 10^7$	$3 \cdot 10^6$	10^5	$8 \cdot 10^4$	500	50	5	1		
1 kleine Bibliothek	10^{12}	$1.5 \cdot 10^{11}$	$2.5 \cdot 10^{10}$	$2 \cdot 10^9$	$2.5 \cdot 10^7$	$2.5 \cdot 10^6$	$2.5 \cdot 10^5$	$5 \cdot 10^9$	1	
1 große Bibliothek	$2.5 \cdot 10^{14}$	$6 \cdot 10^{13}$	10^{13}	$8 \cdot 10^{11}$	10^{10}	10^9	10^8	$2 \cdot 10^7$	400	1

Da aber die Anzahl der Index-Wörter N in Beziehung steht zur Größe des Datenspeichers x, und zwar durch die Relation

$$x = C_2 2^N$$

oder

$$N = C_3 \ln x,$$

ergibt sich

$$\Delta y = C_4 \ln x \,.\, \Delta x \,; \qquad \text{(IL)}$$

(ii) für QA- und CM-Systeme

$$\Delta y = C_5 \Delta x \qquad \text{(QA)}$$
$$\Delta y = C_6 \Delta x \,, \qquad \text{(CM)}$$

wobei die Konstante C_5 die Kosten der Prozessoren ebenso wie die Kosten des Kernspeichers abdecken muß, während C_6 im wesentlichen nur die Kernspeicherkosten enthalten muß.

Die drei obigen Gleichungen können leicht integriert werden und geben dann

$$y = A_o x \log_{10} \frac{x}{c} + A_2 \qquad \text{(IL)}'$$
$$y = B_1 x + B_2 \qquad \text{(QA)}'$$
$$y = C_1 x + C_2 \qquad \text{(CM)}'$$

wobei die Quantitäten A_1 ($= A_o \log_{10} \frac{x}{c}$), B_1 und C_1 die Kosten der entsprechenden Systeme pro Einheit darstellen, und A_2, B_2 und C_2 die Anfangsinvestitionen auch schon für das kleinste System ($x \to 0$) der drei Typen bedeuten.

Die sechs Konstanten sind nach den Erfahrungen mit bereits arbeitenden Systemen grob geschätzt und in Tabelle III zusammengestellt worden. Abbildung 6 zeigt die Kostenfunktionen.

Abbildung 6 macht deutlich, warum Indexiersprachen für kleine Bibliotheken (bis zu 20.000 Bücher oder 100.000 Artikel) die beliebteste Lösung für die Probleme des Bibliothekars sind: sie sind weitaus billiger als alles andere. Kopfschmerzen bekommt er jedoch, sobald sich die Leistungen der Bibliothek auf einen Bestand von einer Million oder mehr Einheiten erstrecken müssen. Sollen

Tabelle III
Geschätzte Installationskostenkonstanten für ein IL-System (A_1, A_2), für ein QA-System (B_1, B_2) und für ein CM-System (C_1, C_2)

	10^3	10^5	10^3	
A_1	280	500	800	
B_1	30			Cents/Einheit
C_1	1			

A_2	10^3	
B_2	$3 \cdot 10^4$	Dollar
C_2	$3 \cdot 10^5$	

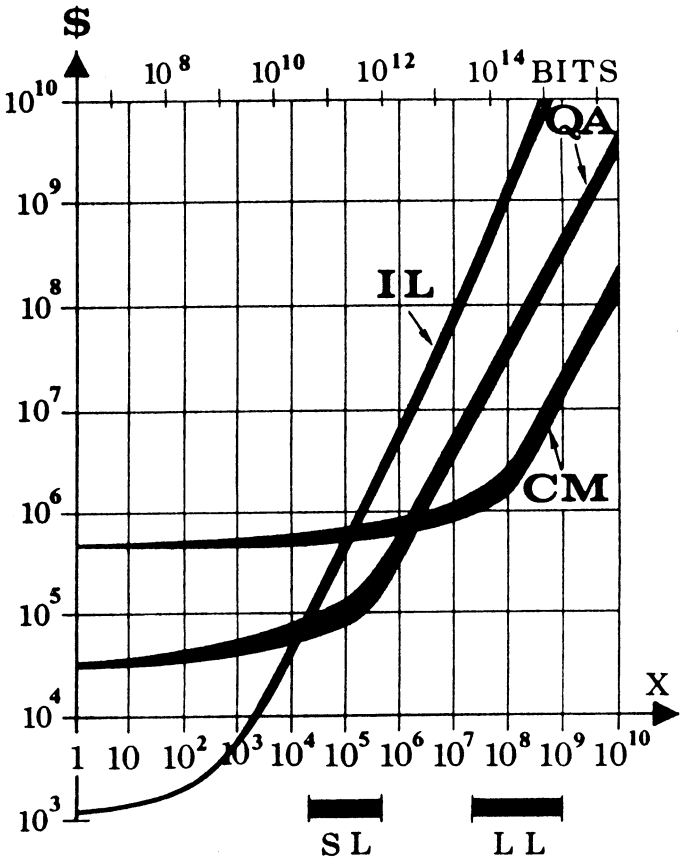


Bild 6 Installationskostenfunktionen für ein IL-System, IL, für ein QA-System, QA, und für ein CM-System, CM. Die Abszisse zeigt die Speicherkapazität an, unten nach der Zahl der Zeitschriftenhefte (oder Buchkapitel), oben in Bits. Die Ordinate gibt die Kosten in US-Dollar an. (SL = Kleine Bibliothek; LL = Große Bibliothek.)

diese noch erbracht werden können, braucht der Bibliothekar Gelder in Höhe von etwa 10 Mio. Dollar. Da er solche Summen aber nie erhalten wird, muß er sich im Bereich der QA- und CM-Systeme nach Angeboten umsehen, die die gewünschte Leistung bieten und viel weniger, etwa um 1 Mio. Dollar, kosten. Wenn sein System aber weiter wächst, wird er gut daran tun, ein CM-System zu installieren, denn die Kosten dafür, etwa 2—3 Mio. Dollar, sind im Vergleich zu den anderen Systemen immer noch ein Pappenstiel.

Mit Abbildung 6 habe ich nicht nur einen Hinweis auf die Investitionsstrategien des Bibliothekars, sondern auch die Antwort auf das vorhin gestellte Problem der Finanzierung der neuartigen Systeme geben wollen, denn gerade die Höhe der Anfangsinvestition für ein CM-System, etwa eine halbe Mio. Dollar, das nicht

mehr leistet als ein IL-System für ein paar tausend Dollar, macht das Finanzierungsproblem zu mehr als einer trivialen Angelegenheit.

Nach dem Versuch einer Beantwortung der Investitionsfrage sollten wir uns nun der Kosten-Leistungs-Frage zuwenden. Ich werde auf diese Pflichtübung aber verzichten, denn es ist nachgewiesen worden (Weston 1971), daß nicht die maschinenorientierten Kosten von Belang sind, sondern die gesellschaftlichen Aufwendungen für den Betrieb und die Instandhaltung solcher Systeme.

Weston hat in einer umfassenden Arbeit (1971) dargelegt, daß die Größe einer Wissenschaftlergemeinschaft – wie sie z. B. durch die Zahl der Mitglieder einer wissenschaftlichen Vereinigung bestimmt werden kann – durch die Anzahl der Dokumente begrenzt wird, die ihre Mitglieder erzeugen und lesen können. Wächst die Gemeinschaft über diese Grenze hinaus, beginnt sie sich in „Spezialistengruppen" aufzusplittern, die ihre eigenen Zeitschriften, Tagungen und Gremien haben und mit der ursprünglichen Gemeinschaft nur mehr den Namen sowie den Präsidenten teilen. Mit anderen Worten, es gibt eine Obergrenze n für den Prozentsatz an Arbeitszeit, den ein Wissenschaftler bereit ist, für die Kommunikation mit seinen Kollegen aufzuwenden, sei es als Produzent, sei es als Rezipient. Nehmen wir an, es gäbe N Wissenschaftler in den USA, von denen jeder (im Durchschnitt) S Dollar verdient, dann beträgt der Preis P für den Austausch von Information durch Bücher und Aufsätze pro Jahr

$P = \eta$ NS \$/Jahr.

Ich überlasse die Ermittlung dieser Zahlen Ihrer Phantasie, wage aber zu behaupten, daß Sie Mühe haben werden, unter einer jährlichen Aufwendung von mehreren Mio. Dollar zu bleiben. Und nimmt man darüber hinaus die Kosten für Bücher und Zeitschriften für längere Zeiträume hinzu, dann ergeben sich astronomische Zahlen. Es ist daher eine von vornherein fehlorientierte Strategie, Investitionen für die Steigerung der Zugriffsleistung zu diesen Trägern potentieller Information zu tätigen: diese Informationsträger sind zu langsam und zu teuer. Die richtige Strategie besteht darin, in die Beschleunigung und Erleichterung der Entwicklung jener neuartigen Träger zu investieren, deren Strukturprinzip in der Erhaltung verbundener semantischer „Nachbarschaften" liegt, und nicht in der *ad-hoc*-Bereitstellung dieser Nachbarschaften mit Hilfe gewisser Artefakte aus einer Ansammlung unverbundener Einheiten, die über mehrere Dokumente verstreut sind.

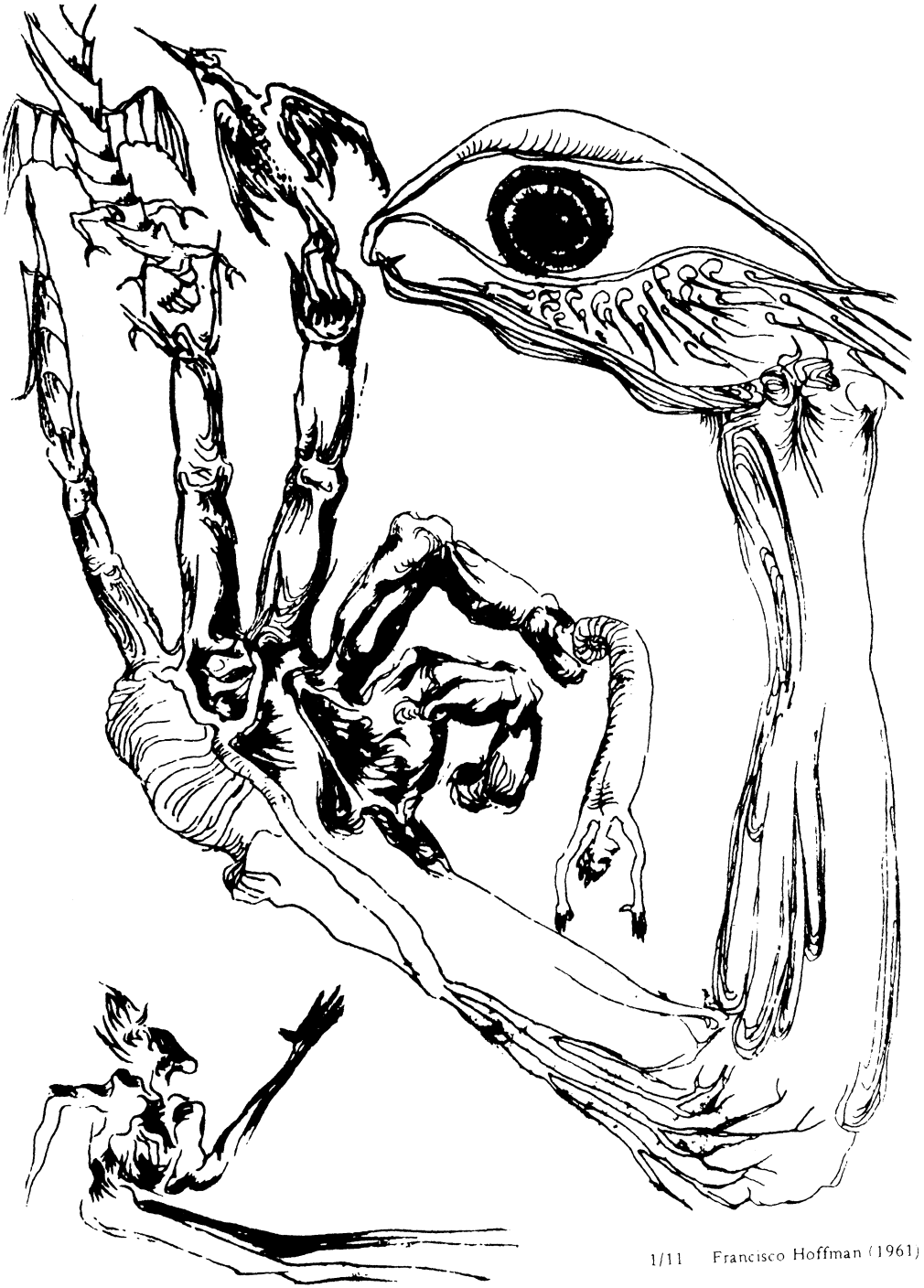
Was ich hier herausstellen möchte, das ist die Beachtung der ehrwürdigen Geschäftsmaxime: Wenn das Produktionsvolumen groß ist, dann sind Investitionen zur Erhöhung der Produktionsleistung gewinnbringend. Da das Volumen, mit dem wir es hier zu tun haben, außerordentlich groß ist, es handelt sich nämlich um das Produkt NS, d. h. die landesweit aufgewendeten Geldbeträge für die Gehälter in Wissenschaft und Technik, muß jede Investition, die die Quantität n, die „Kommunikationsviskositätskonstante", reduziert, und sei es auch nur um ein wenig, beträchtlichen Gewinn abwerfen.

Wenn Computerterminals in allen Universitäten, medizinischen Zentren, Forschungslabors der Industrie usw. mit einem zentral gelegenen vollentwickelten CM-System verbunden wären, dann müßten keine Bücher und Übersichtsartikel mehr geschrieben werden. Die Originalergebnisse und ihre Begründung können di-

rekt in das CM-Datenzentrum eingegeben werden und stehen dann in beliebiger Verbindung oder Beziehung mit anderen Erkenntnissen dem Benutzer zur Verfügung, der derartige Verbindungen und Beziehungen ergründen will, ohne daß er durch den Zwang frustriert wird, die Grenzen von Disziplinen, Zeitschriften, Büchern und Fachbereichen zu überschreiten.

Wieder also sind wir auf das Buch – oder jede gleichartige Form von Dokumentation – als den Engpaß menschlicher Kommunikation gestoßen. Daß ich es wagte, diese Ansicht ausgerechnet Bibliothekaren vorzutragen, kann ich nur durch meine Überzeugung rechtfertigen, daß Sie Ihren Beruf sicherlich nicht allein des Buches wegen gewählt haben, sondern weil Sie anderen Menschen bei der Verwirklichung jener Wünsche helfen wollten, von denen schon Aristoteles sprach: „Alle Menschen sind von Natur aus begierig zu wissen".

Kybernetik einer Erkenntnistheorie



1/11 Francisco Hoffman (1961)

Kybernetik einer Erkenntnistheorie*

Zusammenfassung

Wenn „Epistemologie" nicht als Theorie der Erkenntnis bzw. des Wissens an sich, sondern als Theorie des Erkenntnis- und Wissens*erwerbs* verstanden wird, dann ist – so wird behauptet – der für eine solche Epistemologie angemessene begriffliche Rahmen jener der Kybernetik, denn die Kybernetik ist die einzige wissenschaftliche Disziplin, die eine strenge Behandlung kreis-kausaler Phänomene ermöglicht. Die Prozesse, durch die Erkenntnis bzw. Wissen erworben wird, d.h. die kognitiven Prozesse, werden als algorithmische Rechenprozesse aufgefaßt, die ihrerseits errechnet werden. Dies erfordert die Erörterung von Rechenprozessen, die Rechenprozesse errechnen usw., d.h. die Erörterung von rekursiven Rechenprozessen mit einer Regression von beliebiger Größenordnung.

Aus dieser Perspektive werden die Aktivität des Nervensystems, einige Experimente, die Grundlagen einer künftigen Theorie des Verhaltens und deren ethische Konsequenzen diskutiert.

Als ich zusagte, meinen Vortrag hier auf Deutsch zu halten, hatte ich keine Ahnung, in was für Schwierigkeiten ich geraten würde. In den vergangenen 20 Jahren habe ich wissenschaftlich ausschließlich englisch gedacht und geredet. Viele Begriffe und Forschungsergebnisse wurden bei ihrer Entstehung auf englisch getauft und sträuben sich gegen jede Übersetzung.

Ich habe schließlich, nach vergeblichen Bemühungen um Verdeutschung, beschlossen, sozusagen aus mir selber herauszutreten, das ganze Gedankengewebe wie einen Gobelin zu betrachten und, so gut es geht, Ihnen auf Deutsch zu beschreiben was da abgebildet ist. Sollte ich mich in labyrinthische Satzkonstruktionen verirren, dann ist das keine Affektiertheit, sondern das Ächzen einer verrosteten Maschine.

Vielleicht sind Sie schon beim Lesen meines Vortragstitels gestolpert: „Kybernetik einer Erkenntnistheorie"? Im stillen waren Sie überzeugt, daß ich damit „Eine Erkenntnistheorie der Kybernetik" gemeint hatte. Ursprünglich war das auch so. Aber im Laufe einigen Nachdenkens ist mir klar geworden, daß nicht nur eine Erkenntnistheorie der Kybernetik, sondern jede Erkenntnistheorie, die Anspruch auf Geschlossenheit und Vollständigkeit erhebt, im Grunde eine kybernetische Theorie ist.

Der wesentliche Beitrag der Kybernetik zur Erkenntnistheorie ist die Möglichkeit der Verwandlung eines offenen Systems in ein geschlossenes System, im

---

* Vortrag auf dem 5. Kongreß der Deutschen Gesellschaft für Kybernetik, Sektion Biokybernetik, Nürnberg 28.–30. März 1973.

besonderen Fall das Schließen des linearen, offenen, unendlichen Kausalnexus zu einem geschlossenen, und endlichen, Kausalkreis.

Hier ist vielleicht eine historische Fußnote am Platz. Mehrere Jahre, bevor Norbert Wiener unser Fachgebiet Kybernetik nannte, fanden in New York jährliche Symposien statt, wo ein Kreis von Forschern aus verschiedenen Gebieten, die Wiener nahestanden, zusammenkam und über die gemeinsamen Probleme diskutierte. Das Thema war: „Circular Causal and Feedback Mechanisms in Biological and Social Systems" („Kreiskausal geschlossene und rückgekoppelte Mechanismen in biologischen und sozialen Systemen") (v. Foerster ed. 1949; v. Foerster/Mead/Teuber eds. 1950, 1951, 1953, 1955).

Zunächst hat die Idee des geschlossenen Kausalkreises die angenehme Eigenschaft, daß für eine Wirkung in der Gegenwart die Ursache in der Vergangenheit liegt, wenn man den Kreis an einer Stelle durchschneidet, die Ursache aber in der Zukunft liegt, wenn man ihn an der gegenüberliegenden Stelle schneidet. Der geschlossene Kausalkreis überbrückt also die Kluft zwischen effektiver und finaler Ursache, zwischen Trieb und Zweck.

Zweitens scheint man mit der Schließung der Kausalkette den Vorteil erkauft zu haben, einen Freiheitsgrad losgeworden zu sein: Man braucht sich nicht mehr um die Anfangsbedingungen zu kümmern, denn sie werden ja automatisch von den Endbedingungen geliefert. Das ist zwar in der Tat der Fall, aber einfach ist die Geschichte nicht: Nur ganz bestimmte Zustandsgrößen geben eine Lösung für die Vorgänge im Kreis; das Problem ist ein Eigenwert-Problem geworden.

Erschwerend kommt noch dazu, daß der Verdacht auftaucht, daß die ganze Sache mit dem geschlossenen Kreis eine logische Spitzbüberei sein könnte. Das kennt man ja schon von der Theorie der logischen Urteile; da gibt es den berüchtigten *circulus vitiosus*, den „Teufelskreis": Der Grund wird zur Folge und die Folge zum Grund.

Meine Absicht ist es, den Teufelskreis, den *circulus vitiosus*, nicht nur von jeder schlechten Nachrede zu befreien (Katz 1962), sondern ihn sogar zu der ehrenwerten Position eines *circulus creativus*, eines „schöpferischen Kreises" zu erheben.

Dem will ich zwei Propositionen vorausschicken. In der ersten benütze ich die Ausdrücke „Sensorium" und „Motorium". Unter Sensorium verstehe ich das System der bewußten Sinneswahrnehmungen und unter Motorium das der gewollten Bewegungsabläufe.

Meine erste Proposition:

> „Der Sinn (oder die Bedeutung) der Signale des Sensoriums wird durch das Motorium bestimmt, und der Sinn (oder die Bedeutung) der Signale des Motoriums wird durch das Sensorium bestimmt."

Das heißt, daß *Information* – nicht im informationstheoretischen, sondern im umgangssprachlichen Sinn – ihren Ursprung in diesem *circulus creativus* hat. Be-Deutung hat nur, was ich be-greifen kann.

Meine zweite Proposition bezieht sich auf das Problem einer vollständigen und geschlossenen Theorie der Hirnfunktionen. Sollte einer von uns Sterblichen sich je mit diesem Problem beschäftigen, so wird er zweifellos sein Hirn dazu benützen. Diese Beobachtung liegt meiner zweiten Proposition zugrunde.

Meine zweite Proposition:

„Die Sätze der Physik, die sogenannten ‚Naturgesetze', können von uns geschrieben werden. Die Sätze der Hirnfunktionen oder – noch allgemeiner – die Sätze der Biologie müssen so geschrieben sein, daß das Schreiben dieser Sätze von ihnen abgeleitet werden kann, das heißt: sie müssen sich selber schreiben."

Lassen Sie uns nun zu meinem Thema zurückkehren, nämlich, daß Erkenntnistheorie im wesentlichen eine Kybernetik ist. Wenn ich Er-Kenntnistheorie sage und nicht nur Kenntnistheorie, dann will ich damit betonen, daß die Vorsilbe ER- etwas Schöpferisches, eine Ontogenetik andeutet. Im Griechischen ist das klarer, da taucht die indogermanische Wurzel „GN" sowohl in γίγνομαι („entstehen") wie in γιγνώσκω („Erkennen durch die Sinne") in gleicher Weise auf. Aber neben dem „Kennen durch die Sinne" gibt es im Griechischen auch ein „Kennen durch die Muskeln" ἐπίσταμαι, genau genommen ein handwerkliches Können, ein Verstehen durch das Tun. Die Etymologie hier ist ἐπι- („darüber-") und ἵσταμαι („stehen"), also darüber-stehen, etwa wie im Englischen, wo man allerdings darunter steht („understand"). Warum man, wenn es sich um Erkenntnis handelt, im Deutschen *Ver*-stehen sagt, so als ob man sich den Fuß *ver*-staucht, müssen mir Etymologen erklären, die mehr verstehen als ich.

Wie dem auch sei, wenn ich Er-Kenntnis sage, möchte ich den Prozeß verstehen, durch den Kenntnis erworben wird. Das Zeitwort, das diesen Prozeß beschreibt, wäre dann, so würde man glauben, „erkennen". Leider scheint mir das semantisch zu seicht, denn oft spricht man von erkennen, wenn eigentlich „wiedererkennen" gemeint ist. Wenn es das Wort „Er-wissen" gäbe, so hätte ich es mit Vergnügen gegen „ver-stehen" und „er-kennen" eingetauscht, um genau das zu sagen, was ich sagen möchte[1]).

Zunächst müssen wir den monolithischen Begriff „Erkenntnis", oder den *Prozeß des Erwerbens von Kenntnis* = „Er-Kennen" so paraphrasieren, daß das Wesentliche getroffen wird, uns aber gleichzeitig die Möglichkeit geboten wird, unsere konzeptuellen Werkzeuge verwerten zu können.

Ich schlage daher vor, Er-Kennen durch folgende semantisch analoge Paraphrase zu ersetzen[2]):

Er-Kennen ⟶ Er-Rechnen einer Realität.

In beiden Fällen gebrauche ich die Vorsilbe ER in ihrem ontogenetischen Sinn, das heißt Er-Kennen und Er-Rechnen als einen ständig vor sich gehenden Prozeß und nicht als ein stationäres Resultat.

Hier werden wahrscheinlich einige Skeptiker stutzig. Warum „eine" Realität, warum nicht „die" Realität? könnte man fragen. Hier sind wir doch, der Kyberne-

[1]) Gemäß der freundlichen Beratung eines anwesenden Semantikers bedeutet die Vorsilbe ER- einen all-umfassenden Einschluß, das heißt, daß mein Gebrauch dieser Vorsilbe in ER-kennen und ER-rechnen in der Tat die semantische Tiefe besitzt, die ich auszudrücken hoffte. Bezüglich der Vorsilbe VER- hat derselbe Berater mich aufmerksam gemacht, daß VER- einerseits auf eine Auslöschung hinweist (ver-schwinden, ver-lieren, usw.), andererseits aber mit FÜR- sinnähnlich ist (ver-stehen, für-stehen), meinend vielleicht, daß ein Begriff FÜR das steht, was er bedeuten soll (Symbol).

[2]) Der Pfeil „⟶" soll hier bedeuten „ist interpretierbar als..."

tische Kongreß, die Meistersingerhalle in Nürnberg, das physikalische Universum, als ob es noch eine andere Realität gäbe.

In der Tat trennt ein tiefer epistemologischer Abgrund die beiden Auffassungen, die sich durch die Benützung des bestimmten oder des unbestimmten Artikels voneinander unterscheiden. Der fundamentale Unterschied besteht darin, daß man einmal annimmt, daß unabhängige Beobachtungen die Realität *bestätigen*, zum anderen, daß man annimmt, daß durch *Bezugsetzung* unabhängiger Beobachtungen eine Realität geschaffen wird.

Der erste Fall: Mein visueller Sinn sagt mir, hier steht ein Stehpult. Mein Tastsinn bestätigt das, ebenso der Herr Vorsitzende, Dr. Küpfmüller, würde ich ihn fragen.

Der zweite Fall: Mein visueller Sinn sagt mir, da steht was, mein Tastsinn sagt mir, da steht was; und Dr. Küpfmüller höre ich sagen: „Da steht ein Stehpult". Die Bezugsetzung dieser unabhängigen Beobachtungen erlaubt mir, etwas damit anzufangen, z.B. zu sagen: „Hier steht ein Stehpult".

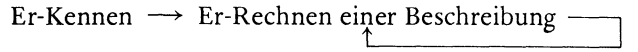
Im ersten Fall postulieren wir Bestätigung, Konfirmation, als unser Arbeitsprinzip, im zweiten Fall Bezugsetzung oder Korrelation. Ich habe jedenfalls die Absicht, den unbestimmten Artikel zu benützen, denn er gibt den allgemeinen Fall. Gewiß ist *die* Realität ein spezieller Fall *einer* Realität.

Kaum haben wir den unbestimmten Artikel zur Ruhe gelegt, taucht eine neue Frage auf: Was soll denn das „Errechnen"? Ich könnte, so meint man, doch nicht allen Ernstes behaupten, daß ich das Stehpult, oder meine Armbanduhr, oder den Andromeda-Nebel „errechne"? Allenfalls könnte man sagen, daß eine „Beschreibung einer Realität" errechnet wird, denn mit meinen verbalen Hinweisen „Stehpult", „Armbanduhr", „Andromeda-Nebel" habe ich gerade demonstriert, daß gewisse Bewegungsabläufe meines Körpers, die mit gewissen Zisch- und Grunzlauten verbunden waren, bei den Zuhörern die Interpretation einer Beschreibung zuließen.

Die folgende Korrektur ist daher angebracht:

Er-Kennen ⟶ Er-Rechnen einer Beschreibung einer Realität.

Die Neurophysiologen können hier einwenden, daß die Eingangssignale erst viele Modifikationsstufen durchlaufen, ehe wir ein verbales Ausgangssignal erhalten. Zunächst entsteht auf der Netzhaut eine zweidimensionale Projektion der Außenwelt, die man als „Beschreibung erster Ordnung" bezeichnen kann. Das angeschlossene post-retinale Netzwerk liefert dann den Ganglienzellen eine modifizierte Beschreibung dieser Beschreibung, also eine „Beschreibung zweiter Ordnung". Und so geht das weiter über die verschiedenen Stationen der Errechnung zu Beschreibungen höherer und höherer Ordnungen. Wir können daher die zweite Fassung meiner Proposition so erweitern:



Die Paraphrasierung hat zwei Vorteile. Der Stein des Anstoßes, „die" oder „eine" Realität, ist verschwunden, denn hier ist von Realität schon nicht mehr die Rede. Zweitens können wir von der Einsicht Gebrauch machen, daß das Errechnen einer Beschreibung ja nichts anderes ist als eine Errechnung. Damit kom-

men wir zu einer endgültigen Paraphrasierung des unaufhörlich sich neu einleitenden Vorgang des Er-wissens oder Er-kennens, nämlich:

Ich deute also Erkenntnis oder den Prozeß der Erwerbung von Kenntnis als rekursives Errechnen.

An dieser Stelle stehe ich an einem Gabelpunkt: Einerseits könnte ich jetzt über Eigenschaften von rekursiven Funktionen sprechen und von den Eigenschaften von Maschinen, die rekursive Funktionen rechnen. In Maschinensprache würde ich dann von Kaskaden von Compiler-Sprachen sprechen und über die Theorie der Meta-Programme. In diesem Fall ist natürlich der Begriff der Turingmaschine das ideale konzeptuelle Werkzeug. Zum Beispiel wird es vollkommen klar, daß die Struktur der Quadrupel einer Turingmaschine, die die Kode-Nummer einer anderen Maschine oder ihre eigene berechnen, mit der Struktur der errechneten Bandbeschreibungen nicht verwechselt werden kann (Löfgren 1967). Oder neurophysiologisch ausgedrückt: Um „Stehpult" zu sagen, oder zu wissen, daß hier ein Stehpult steht, muß ich weder im Gehirn die Buchstaben STEHPULT stehen haben, noch braucht eine winzige Repräsentation eines Stehpults irgendwo in mir zu sitzen. Ich brauche aber eine Struktur, die mir die verschiedenen Manifestationen einer Beschreibung errechnet. Aber das gehört alles in die Sitzung über „Künstliche Intelligenz", und hier sind wir in der Sitzung „Biokybernetik des Zentralnervensystems". So finde ich es besser angebracht, über die Auswirkung meiner Proposition von vorher auf die Deutung der Vorgänge im Zentralnervensystem zu sprechen.

Zunächst wollen wir noch einmal feststellen, mit was für einem ungeheuren Problem wir es zu tun haben. Zu diesem Zweck rufen wir uns den Satz der undifferenzierten Kodierung ins Gedächtnis zurück.

Der Satz der undifferenzierten Kodierung:

„Die Erregungszustände einer Nervenzelle kodieren *nicht* die Natur der Erregungsursache. (Kodiert wird nur: ‚so und so viel an dieser Stelle meines Körpers', aber nicht ‚was'.)"

Ein Beispiel: Ein Stäbchen in der Netzhaut absorbiert in einem gewissen Moment einen Photonenstrom von so und so vielen Photonen pro Sekunde. Es entwickelt dabei ein elektrochemisches Potential, das eine Funktion des Stromes ist, das heißt, daß das „Wieviel" kodiert wird; aber die Signale, zu denen dieses Potential schließlich Anlaß gibt, geben weder einen Aufschluß darüber, daß Photonen die Reizursache waren, und schon gar nicht, aus was für einem Frequenzgemisch der Photonenstrom bestand. Ganz genau dasselbe ist für alle anderen Sinneszellen der Fall. Nehmen Sie die Haarzellen in der Cochlea oder die Meissnerschen Tastkörperchen oder die Geschmackspapillen, oder was immer für Zellen Sie wollen, in keiner wird die *Qualität* der Erregungsursache kodiert, nur die *Quantität* der Erregung. Und in der Tat, „da draußen" gibt es ja kein Licht und keine Farben, da gibt es elektromagnetische Wellen; „da draußen" gibt es keine Laute und keine Musik, da gibt es longitudinale periodische Druckvorgänge; „da draußen" gibt es keine Hitze oder Kälte, da gibt es eine höhere oder niedrigere mittlere kinetische Molekularenergie, und so weiter – und ganz bestimmt: „da draußen" gibt es keinen Schmerz.

Die fundamentale Frage heißt dann: Wieso erleben wir die Welt in ihrer überwältigenden Mannigfaltigkeit, wenn als Eingangsdatum uns lediglich zur Verfügung steht: erstens die Reizintensität; zweitens die Koordinaten der Reizquelle, das heißt Reizung an einer bestimmten Stelle meines Körpers?

Nachdem die Qualitäten der Sinneseindrücke nicht im Empfangsapparat kodiert sind, ist es klar, daß das Zentralnervensystem so organisiert ist, daß es diese Qualitäten aus diesen kümmerlichen Eingängen errechnet.

Über diese Operationen wissen wir – oder zumindest ich – im Detail sehr wenig. Ich freue mich daher schon auf mehrere Vorträge, die sich mit diesem Problem befassen werden, besonders auf den von Donald MacKay, Horst Mittelstaedt und Hans Lukas Teuber, die sich mit dem berühmten Re-Afferenzprinzip befassen werden.

Ich werde mich damit begnügen müssen, nur ein paar Hinweise über die Art dieser Operationen zu geben.

Lassen Sie mich zunächst mit einem pädagogischen Kunstgriff das Problem der Errechnung perzeptueller Mannigfaltigkeiten so formulieren, daß durch diese neue Perspektive das Problem vom perzipierenden Organismus her gesehen wird und nicht – wie üblich – vom Standpunkt eines Beobachters, der, durch seine eigene Perzeption verführt, immer schon glaubt zu wissen, wie es „da draußen" aussieht, und dann mit Mikropipetten im Nervensystem des Organismus festzustellen sucht, wie „das da draußen" denn „da drinnen" aussieht. Epistemologisch gesehen ist das aber eine Form von Mogeln, denn der Beobachter schielt sozusagen auf die Seite nach den „Antworten" (sein eigenes Weltbild), die er dann mit irgendwelchen Zellerregungszuständen vergleicht, aus denen allein aber der Organismus sein Weltbild zusammenstückeln muß. Wie er das tut, das ist aber das Problem (v. Foerster 1970a).

Es ist klar, daß jeder Organismus ein von einer geschlossenen Oberfläche begrenztes endliches Volumen hat, das von einem verzweigten Röhrensystem durchzogen ist. Dieses durchbricht an mehreren, sagen wir s, Stellen die Oberfläche. Ontogenetisch gesehen wird die Oberfläche vom Ektoderm bestimmt und enthält alle sensiblen Endorgane, während das Innere vom Endoderm abzuleiten ist. Topologisch gesehen stellt eine solche Oberfläche eine orientierbare zweidimensionale Mannigfaltigkeit von der Ordnung $p = (s + t)/2$ dar, wo t die Anzahl der T-Verbindungen des Röhrensystems bedeutet. Nach einem bekannten Satz der Topologie ist aber jede geschlossene und orientierbare Oberfläche von endlicher Ordnung metrisierbar, das heißt wir können auf die Oberfläche eines jeden Organismus ein geodätisches Koordinatensystem legen, das in unmittelbarer Nachbarschaft eines jeden Punktes euklidisch ist. Wir wollen dieses Koordinatensystem das „autologische" nennen und die beiden Werte ξ_1 und ξ_2, die einen Oberflächenpunkt ein-eindeutig definieren, kurz mit ξ bezeichnen. Damit ist erreicht, daß jeder sensiblen Zelle ein-eindeutig ein Koordinatenpaar zugewiesen ist, das sich ausschließlich auf den Organismus bezieht.

Nach einem ebenso wohlbekannten Satz der Topologie ist aber jede orientierbare Oberfläche der Ordnung p mit einer Kugeloberfläche gleicher Ordnung identisch, das heißt wir können die sensible Oberfläche eines jeden Organismus auf eine Kugel abbilden – die „repräsentative Einheitskugel" (Bild 1), so daß jeder sensiblen Zelle des Organismus ein Punkt auf der Einheitskugel entspricht

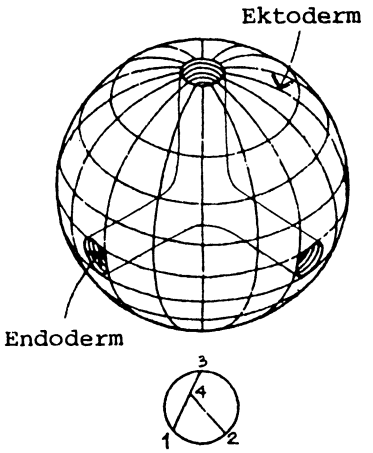


Bild 1
Geschlossene orientierbare Oberfläche zweiter Ordnung.
(s = 3; t = 1; p = (s + t)/2 = 4/2 = 2)

und umgekehrt. Es läßt sich leicht zeigen, daß dieselben Überlegungen auch auf das Innere angewandt werden können.

Es ist klar, daß die repräsentative Einheitskugel invariant bleibt gegenüber allen Deformierungen und Bewegungen des Organismus, das heißt, die autologischen Koordinaten sind absolute Invarianten.

Jedoch von einem Beobachter aus gesehen ist der Organismus in ein euklidisches Koordinatensystem (heterologisches Koordinatensystem x_1, x_2, $x_3 \rightarrow x$) eingebettet, und er mag die Zuordnung $\xi = F(x)$ als die „Formfunktion", F, oder $\xi = \Phi(x,t)$ als die Verhaltensfunktion, Φ, bezeichnen ($t \rightarrow$ Zeit)[3]). In Bild 2a ist diese Situation für eine fischförmige Kreatur dargestellt. Die autologischen Koordinaten ξ des Fisches sind hier den heterologischen Koordinaten x überlagert eingezeichnet.

Der pädagogische Kunstgriff, von dem ich vorhin sprach, besteht nun darin, die Koordinatensysteme so zu transformieren, daß die autologischen Koordinaten überall euklidisch werden, das heißt, die Welt des Fisches auf die repräsentative Einheitskugel abzubilden (Bild 2b, ξ sind Polarkoordinaten). Damit aber wird die Welt „da draußen" nicht mehr euklidisch, wie aus der Divergenz und Konvergenz der vertikalen x-Koordinaten deutlich zu sehen ist. Eine Bewegung des Fisches, die vom Beobachter wie in Bild 3a gesehen wird, bedeutet dann für den Fisch, daß er durch die Anspannung seiner Muskeln die Metrik seiner Umwelt verändert hat (siehe Bild 3b); er selbst ist natürlich immer derselbe geblieben (Invarianz der Selbst-Referenz), wie das ja die repräsentative Einheitskugel zum Ausdruck bringt.

Man beachte die Sehlinie Auge–Schwanz, die vom Beobachter als Gerade interpretiert wird (Bild 3a), vom Fisch aber, der über die gekrümmte Geodätische

---

[3]) Für nicht deformierbare (exo-, endo-skelettige) Lebewesen ist natürlich $F(x) = \Phi(x,t) = 0$, und außerdem gilt allgemein wegen (angenäherter) Inkompressibilität:

$$\oint_F dV = \oint_\Phi dV = V(t) = \text{const.}$$

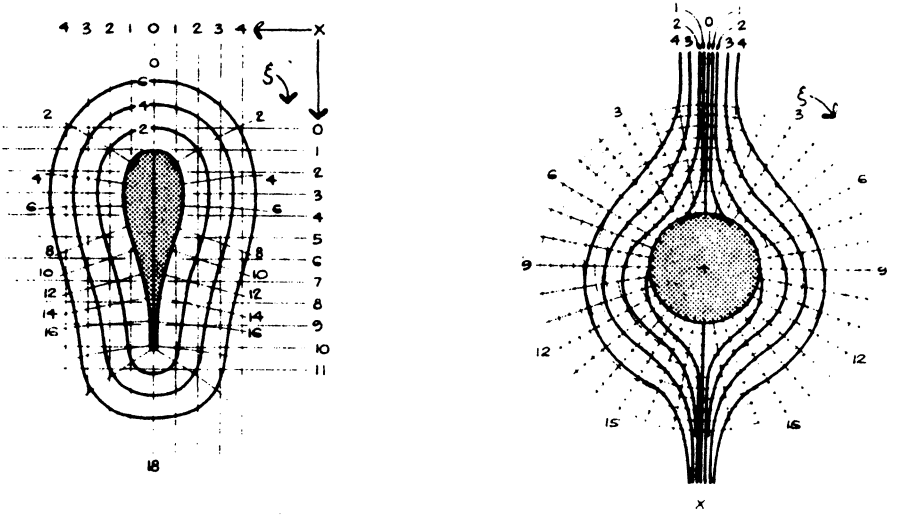


Bild 2a Geodätische Koordinaten des subjektiven Koordinatensystems, ξ, eines Organismus in Ruhe, der in eine Umwelt mit euklidischer Metrik eingebettet ist.

Bild 2b Euklidische (polare) Koordinaten des subjektiven Koordinatensystems, bezogen auf die repräsentative Einheitskugel, die in eine Umwelt mit einer dem Ruhezustand des Organismus entsprechenden nicht-euklidischen Geometrie, x, eingebettet ist.

in Bild 3b seinen Schwanz sieht, als solche erst errechnet werden muß. Für diese Errechnung stehen ihm aber nur die Erregungswerte $E(\xi)$ längs seiner Oberfläche zur Verfügung. Obwohl ich hier nicht in aller Strenge zeigen möchte, daß dieses Datum unzureichend ist, um die Beschaffenheit seiner „Umwelt" zu errechnen[4]), so hoffe ich doch, daß aus diesen Bemerkungen deutlicher als sonst hervorgeht, daß nur die Bezugsetzung der Motoraktivität des Organismus mit den so veränderten Erregungen seiner Sinnesorgane es ihm überhaupt erst möglich macht, diese Erregungen eindeutig zu interpretieren. Hier taucht wieder, in etwas veränderter Form, das Eigenwertproblem auf, das ich schon früher erwähnte und von dem später noch die Rede sein wird.

Wir sind zur Zeit damit beschäftigt, die Proposition, daß das Motorium die Interpretation für das Sensorium liefert und daß das Sensorium die Interpretation für das Motorium liefert, ein für allemal experimentell zu etablieren.

Es gibt außerordentlich interessante Experimente mit Kleinstkindern, in denen man zeigen kann, daß das Erwerben perzeptueller Mannigfaltigkeiten direkt mit der Manipulation gewisser geeigneter Objekte im Zusammenhang steht (Piaget/Inhelder 1956; Bower 1971; Witz 1972). Aber leider kann man mit Kleinstkindern nicht reden; oder richtiger, man kann mit ihnen reden, aber man

[4]) Nur für elektrische Aale scheint dies in allen Fällen möglich zu sein, denn die Laplacesche Gleichung ist durch den Randwert längs einer gegebenen Oberfläche eindeutig bestimmt.

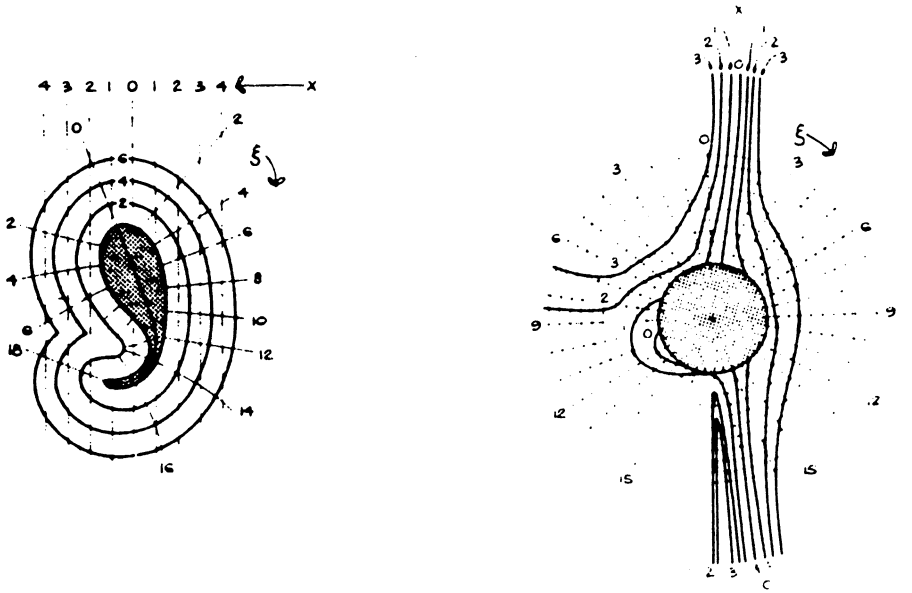


Bild 3a Geodätische Koordinaten des subjektiven Koordinatensystems, ξ, eines Organismus in Bewegung, der in eine Umwelt mit euklidischer Metrik eingebettet ist.

Bild 3b Euklidische (polare) Koordinaten des subjektiven Koordinatensystems bezogen auf die repräsentative Einheitskugel, die in eine Umwelt mit einer dem Bewegungszustand des Organismus entsprechenden nicht-euklidischen Geometrie, x, eingebettet ist.

versteht ihre Antworten nicht. Mit Erwachsenen besteht die Schwierigkeit, daß das senso-motorische System schon so gut integriert ist, daß man schwer trennen kann, was schon früher erlernt wurde und nun in die Experimentalsituation hinüber getragen wird, außer man geht in eine völlig neue „Dimension", deren Zugang unseren früheren Erfahrungen prinzipiell verschlossen blieb.

Diese Dimension wäre zum Beispiel die vierte räumliche Dimension. Wir alle haben einmal früher oder später in unserem Leben die bittere Erfahrung gemacht, daß es unglaublich schwierig, ja sogar unmöglich ist, sich in die vierte Dimension hineinzuzwängen. Da steht an jeder Stelle unseres drei-dimensionalen Raumes eine offene Tür zum Eintritt in die vierte Dimension, aber wie wir uns auch strecken und verrenken, wir bleiben in den allzu bekannten drei Dimensionen stecken.

Wir (im Biological Computer Laboratory) haben uns daher gefragt: Würde der Satz vom Ursprung der Erkenntnis in der Wechselwirkung von Gnosis und Episteme, Sensorium und Motorium, nicht ungeheuer an Plausibilität gewinnen, wenn wir zeigen könnten, daß die vierte Dimension wesentlich *anschaulicher* wird, wenn sie auch *begriffen* werden kann, das heißt, wenn man in die vierte Dimension „hineingreifen" und dort mit vier-dimensionalen Objekten „hantieren" könnte (Arnold 1971; 1972).

Dank der freundlichen Rechenmaschinen, die keine Dimensionalität kennen, läßt sich dieses Vorhaben durchführen, indem man der Versuchsperson (VP)

Manipulatoren zur Verfügung stellt, die „on-line" mit einem schnellen Rechner gekoppelt sind. Gemäß den Stellungen der Manipulatoren schreibt dieser Rechner zwei zugehörige zweidimensionale Projektionen eines vierdimensionalen Körpers auf den Schirm einer Kathodenstrahlröhre, die, von der VP durch ein Stereoskop betrachtet, als drei-dimensionale im Raum schwebende Projektionen dieses Körpers gesehen werden.

Die Konstruktion von Projektionen n-dimensionaler Gebilde (n = 1, 2, ...) auf andere Dimensionen (m < n) läßt sich ohne Schwierigkeit mit einer rekursiven Vorschrift entwickeln, die angibt, wie das (n + 1)-dimensionale Gebilde aus dem analogen n-dimensionalen Gebilde zusammengefügt wird. In den Bildern 4 bis 6 ist dargestellt, wie zum Beispiel ein vierdimensionaler „Würfel" (Tesseract) schrittweise aus dem eindimensionalen „Würfel" (Strecke) entwickelt wird. Die rekursive Vorschrift heißt:

(i) Füge an die 2n Begrenzungen des n-dimensionalen „Grundwürfels" je einen n-dimensionalen Würfel, und füge schließlich an einen von diesen einen „Schlußwürfel".

(ii) Falte die so angefügten Würfel um die 2n Begrenzungen des Grundwürfels in die nächst höhere Dimension, bis sie auf den Begrenzungen senkrecht stehen (das heißt, bis ihre Normalprojektionen verschwinden).

(iii) Falte schließlich den Schlußwürfel und seine gefügte Begrenzung, so daß alle ungefügten Begrenzungen sich fügen.

In den Bildern 4a und 4b ist die Vorschrift (i) für die Erzeugung eines zweidimensionalen Würfels (Quadrat) aus vier eindimensionalen Würfeln (Strecken)

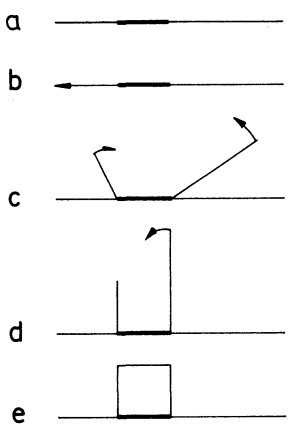


Bild 4 Schrittweiser Aufbau eines 2-D Würfels (Quadrat) durch Faltung von drei 1-D Würfeln (Strecken) aus dem Quellraum (Gerade) um ihre Begrenzungen (Punkte).

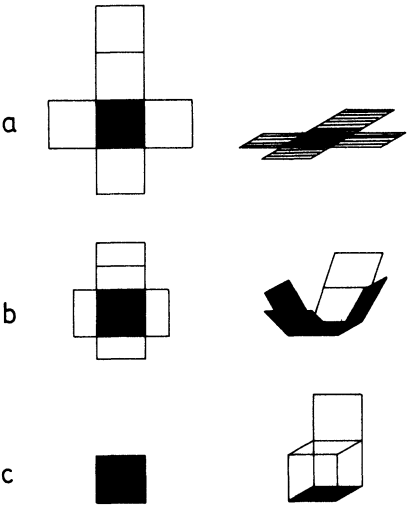


Bild 5 Schrittweiser Aufbau eines 3-D Würfels (Würfel) durch Faltung von fünf 2-D Würfeln (Quadraten) aus dem Quellraum (Ebene) um ihre Begrenzungen (Kanten).

dargestellt; 4c und 4d deuten Vorschrift (ii) an, und 4d und 4e Vorschrift (iii). Für die Erzeugung der drei- und vierdimensionalen Würfel sind die drei Vorschriften in den drei Schritten der Bilder 5 und 6 skizziert. Es ist klar, daß Normalprojektionen der so entstandenen Gebilde in die Entstehungsräume nichts Neues bieten (die Normalprojektion eines Würfels in die Ebene ist ein Quadrat). Erst wenn das Gebilde schräg projiziert wird, wird seine höhere Mannigfaltigkeit offenbar, wie aus Bild 7 hervorgeht, in dem ein mit seinem Grundwürfel (Base) parallel zu unserem Raum stehender Tesseract erst schräg in die dritte und dann normal in die zweite Dimension der Papierebene projiziert ist.

Die Bilder 8a, 8b und 8c geben die stereoskopischen Paare je eines Tesseracts, eines händigen Blocks von einem Soma-Würfel (ein dreidimensionales Gebilde, das Händigkeit unter einer vierdimensionalen Rotation wechselt), und eine Repräsentation einer Klein-Flasche durch zwei Mobiusbänder, bei denen entsprechende Punkte in der vierten Dimension verbunden sind[5]).

In der Versuchsanordnung kann nun entweder der Versuchsleiter (VL), oder die VP selbst die auf den Schirm geschriebenen Projektionen über zwei Manipulatoren mit je drei Freiheitsgraden kontrollieren, wobei die rechte Hand

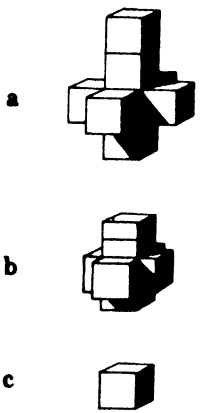


Bild 6

Schrittweiser Aufbau eines 4-D Würfels (Tesseract) durch Faltung von sieben 3-D Würfeln (Würfeln) aus dem Quellraum (Raum) um ihre Begrenzungen (Seiten. (Vgl. 6a mit Salvador Dalis Gemälde im Metropolitan Museum, New York: „Crucifixation (Corpus Hypercubus)". Reproduziert, zum Beispiel, in Descharnes 1962.

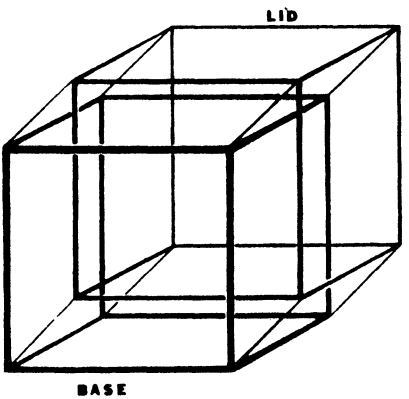


Bild 7

Schrägprojektion eines mit seinem Grundwürfel (BASE) parallel zum Raum des Beobachters stehenden Tesseract (4-D Würfel). (Abschlußwürfel = LID)

[5]) Die Figuren erscheinen ohne Stereoskop dreidimensional, wenn man sie mit gekreuzten Augenachsen fixiert. Fixiere einen zwischen Papierebene und Auge gehaltenen Zeigefinger bis das stereoskopische Paar im Hintergrund in ein einziges Bild verschmilzt. Übertrage Fixierpunkt von Finger zu Papier, ohne die Augenachsen zu verändern. Bei Betrachtung durch ein Stereoskop vertausche links mit rechts (Arnold 1972).

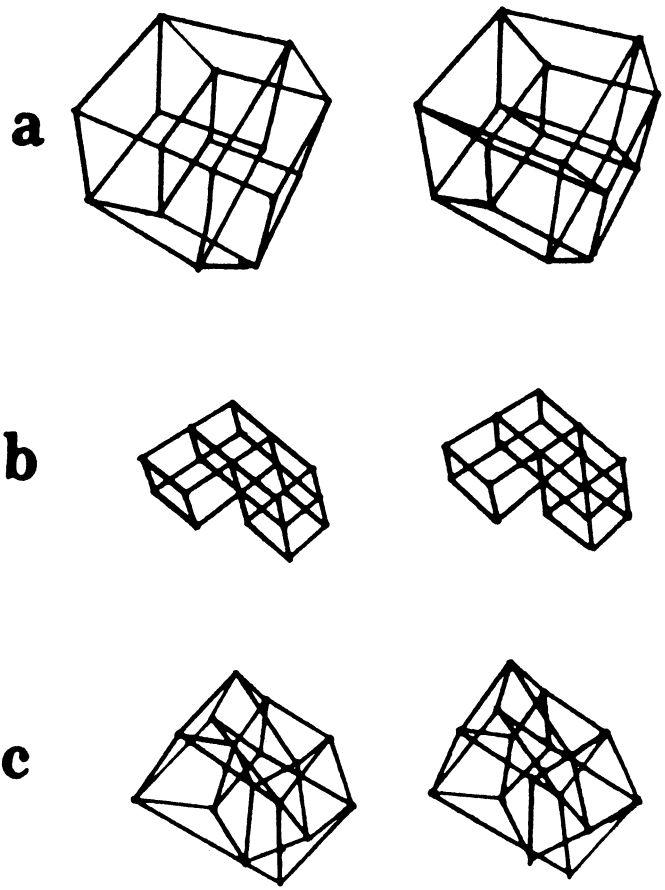


Bild 8 Normalprojektionen von schräg zum Raum des Beobachters stehenden (a) Tesseract; (b) Soma-Würfel; (c) Klein-Flasche: (für stereoskopische Ansicht, siehe Fußnote5).

die drei Rotationen in der xy-, xz-, yz-Ebene unseres Raumes isomorph steuert, die linke Hand aber die drei in der vierten Dimension, nämlich wx, wy, wz. Obwohl die Versuche noch nicht abgeschlossen sind, kann ich berichten, daß die Einsicht, es handle sich bei diesen sich in merkwürdiger Weise verändernden Gebilden um nichts anderes als die Projektionen ein und desselben Objektes – einer „erzeugenden Invariante" –, bei den meisten naiven VP durch ein lautes und begeistertes „AHA!" in etwa 20 bis 40 Minuten mitgeteilt wird, wenn sie die Manipulatoren selber benützen dürfen. Ein ähnlicher Ausruf kommt, wenn überhaupt, von einer gleichen Kategorie von VP, erst nach vier- bis achtstündigen Sitzungen, in denen der VL mit den Manipulatoren hantiert und geduldig immer wieder die geometrische Situation erklärt, wenn gefragt.

Wir hoffen, durch Weiterentwicklung der Apparatur, bei der der Eingriff in die vierte Dimension durch isometrische Kontraktionen der Hals-, Arm- oder Oberschenkelmuskeln bewerkstelligt wird – also durch motorische Beteiligung,

die *keine* 3D-, sondern nur 4D-Konsequenzen zeigt – die Bedeutung der motorisch-sensorischen Wechselwirkung für das Erkennen noch deutlicher zeigen zu können.

Als einen weiteren Hinweis, wie im Zusammenhang mit dem Zentralnervensystem der Begriff „rekursives Errechnen" aufgefaßt werden kann, will ich noch eine andere Beobachtung als Satz zitieren:

Das neurologische (raumzeitliche) Nahwirkungsgesetz:

„Der Erregungszustand einer Nervenzelle ist ausschließlich bedingt durch die (elektro-chemischen) Zustandsgrößen in ihrer unmittelbaren Nachbarschaft (Mikro-Umwelt) und durch ihren (unmittelbar) vorhergehenden eigenen Erregungszustand: Es gibt keine neurologische Fernwirkung."

Es ist ja klar, daß wir nicht auf den Tisch da drüben reagieren, sondern auf die Erregungszustände unserer Stäbchen und Zäpfchen und die unserer Propriozeptoren, die, nach gewissen Operationen im Zentralnervensystem, es uns ermöglichen, auf „den Tisch da drüben" Bezug zu nehmen. Das klingt zwar trivial, dient aber dazu, eine Transformation der Betrachtung zu erleichtern, die, ähnlich wie in der Physik, die Aussage „der Mond wird von der Erde angezogen" (Fernwirkungstheorie) in die Aussage verwandelt, „der Mond bewegt sich im Gravitationsfeld der Erde" (Nahwirkungstheorie). Eine unmittelbare Folge dieser Verschiebung in der Betrachtungsweise ist, daß der übliche Unterschied zwischen „sensiblen" und „schaltenden" Neuronen – oder, wenn Sie wollen, zwischen peripherem und zentralem Nervensystem – verschwindet, denn jede Nervenzelle ist jetzt als „sensible" Zelle aufzufassen, die spezifisch auf ihre Mikro-Umwelt reagiert. Da aber die Totalität der Mikro-Umwelten aller Neuronen eines Organismus seine „Gesamtumwelt" darstellt, ist es klar, daß es nur einem außenstehenden Beobachter vergönnt ist, zwischen einer „äußeren" und einer „inneren" Umwelt eines Organismus zu unterscheiden. Das ist ein Privileg, das dem Organismus selber versagt ist, denn er kennt nur *eine* Umwelt: die, die er erlebt; er kann ja, zum Beispiel, nicht zwischen halluzinatorischen und nicht-halluzinatorischen Erlebniszuständen unterscheiden.

Behalten wir für einen Augenblick die eben erwähnte Unterscheidung des Beobachters bei, dann können wir den Einfluß der beiden Umwelten auf den nervösen Zustand des Organismus im Allergröbsten abschätzen. Ich fasse jede sensible Zelle als einen „Reizpunkt" auf, der mit der äußeren Umwelt („Außenwelt") gekoppelt ist, und jeden synaptischen Dorn einer Zelle des Zentralnervensystems als einen bevorzugten Reizpunkt der „Innenwelt", dessen Mikro-Umwelt durch die chemische Zusammensetzung der Überträgersubstanzen im zugehörigen synaptischen Spalt und durch den elektrischen Erregerzustand des afferenten Axons bestimmt ist. Wenn man als relatives Maß der Wirkungen der Innenwelt und Außenwelt auf den nervösen Zustand des Organismus das Verhältnis der Anzahl der inneren und äußeren Reizpunkte nimmt, dann erhalten wir mit etwa $2 \cdot 10^8$ äußeren und $2 \cdot 10^{13}$ inneren Reizpunkten[6]) eine 100 000-fach höhere Empfindlichkeit des Nervensystems gegenüber Veränderungen der Innenwelt als denen der Außenwelt. Ob sich das Nervensystem diese außerordentliche Sensibilität für innere Veränderungen leisten kann, weil die thermischen und hormonalen Parameter inner-

[6]) Die $2 \cdot 10^{10}$ Neurone des menschlichen Zentralnervensystems haben im Durchschnitt etwa 10^3 Synapsen.

halb der Schädeldecke so unglaublich konstant gehalten werden, oder ob durch diese Konstanz synaptische Proliferation gefördert wird, gehört zu der Kategorie der Fragen: Was war zuerst da, Huhn oder Ei? Jedenfalls mag diese Beobachtung für Neuropharmakologen und Psychiater ein Hinweis sein, daß sie es mit einem höchst empfindlichen System zu tun haben, das möglicherweise schon bei minutiösen Änderungen im Stoffwechsel merkbare Veränderungen in seinem gesamten *modus operandi* zeigen kann, und dem Informatiker mag es sagen, daß er es hier mit einem Rechner zu tun hat, dessen Programmstruktur durch seine Aktivität modifizierbar ist (v. Foerster 1970b).

In Bild 9 ist dem soeben entworfenen Bild der Organisation des Nervensystems schematisch Rechnung getragen. Die schwarzen Quadrate symbolisieren Nervenbündel, die über die Zwischenräume – ein Kollektiv synaptischer Spalte – auf das folgende Bündel wirken können. Der Signalfluß längs der Bündel läuft von links nach rechts, beginnt mit der sensiblen Oberfläche und terminiert in der motorischen Oberfläche, deren Veränderungen über die Außenwelt – den „motorisch-sensorischen synaptischen Spalt" – auf die sensorische Oberfläche rückgekoppelt werden und so den Signalstrom über einen Kreis schließen. Ein zweiter Signalkreis beginnt an der unteren Begrenzung, die den Kontakt des Zentralnervensystems mit der Neurohypophyse schematisiert, die, analog der motorischen Oberfläche, über das vaskuläre System – den „endokrin-operationellen synaptischen Spalt" – die Mikro-Umwelt aller Synapsen kontrolliert (angedeutet durch die feinen Veräderungen in den Zwischenräumen).

Interpretiert man in diesem Schema die Kantenlänge eines Quadrates mit der Anzahl der Reizpunkte im zugehörigen Nervenbündel, dann hätten wir das ganze System aus $10^5 \times 10^5$ Quadraten bestehend skizzieren müssen, um dem vorher erwähnten Übergewicht der inneren gegenüber der äußeren sensiblen Oberfläche Genüge zu tun. Ein Quadrat müßte dann durch einen Punkt mit einem Durchmesser von etwa 1μ dargestellt werden, sollte das Bild nicht größer werden als das hier wiedergegebene.

Um diesem funktionellen Schema auch geometrisch Rechnung zu tragen, können wir die rechtwinklig zueinander fließenden Signalkreise durch Wicklung um eine vertikale und horizontale Achse schließen. Eine ebene Figur, die nach zwei rechtwinkligen Achsen gewickelt wird, ist aber ein Torus. Bild 10 zeigt eine

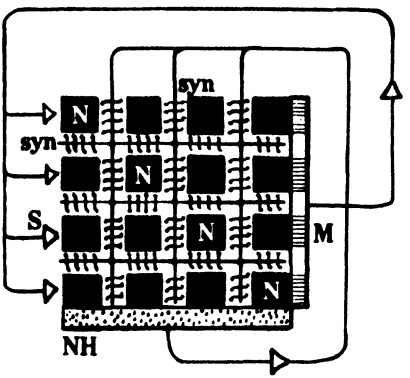


Bild 9
Nervöser Signalfluß von der sensiblen Oberfläche (linke Begrenzung, S), über Nervenbündel (schwarze Quadrate, N) und synaptische Spalte (Zwischenräume, syn) zu Muskelfasern (rechte Begrenzung, M) einerseits, deren Aktivität die Reizverteilung der sensiblen Oberfläche verändert, und Neurohypophyse (untere Begrenzung, NH) andererseits, deren Aktivität die Zusammensetzung der Steroide in den synaptischen Spalten und damit die Funktionsverteilung aller Nervenbündel moduliert.

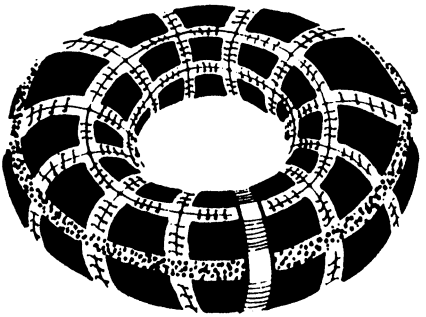


Bild 10
Doppelte Schließung der nervösen und hormonalen Kausalkette.
Horizontale punktierte Naht (Äquator): Neurohypophyse.
Vertikale gestrichelte Naht (Meridian): motorisch-sensorischer „synaptischer Spalt".

vollständige Schematisierung des Gedankens der doppelten Schließung des Signalstroms. Die vordere Naht entspricht dem motorisch-sensorischen synaptischen Spalt, die horizontale Naht der Neurohypophyse.

Dieses Minimalschema der Ur-Organisation eines nervösen Lebewesens hilft vielleicht auch das Problem zu sehen, das entsteht, wenn wir die Vorgänge des Er-Rechnens einer Realität ohne die Zuhilfenahme eines Beobachters, der vorgibt, zwei Seiten zu kennen, lediglich auf Grund rekursiver Rechenoperationen innerhalb des Organismus ableiten wollen, kurz gesagt: wenn wir eine geschlossene Theorie des Beobachters entwickeln und für diese Theorie nicht sofort wieder einen Beobachter zweiter Ordung – usw. – zu Hilfe rufen wollen (Maturana 1970a; 1970b).

Als Leitfaden für die Erforschung dieses Problems möge der folgende Satz dienen:

Das Postulat der epistemischen Homöostase:

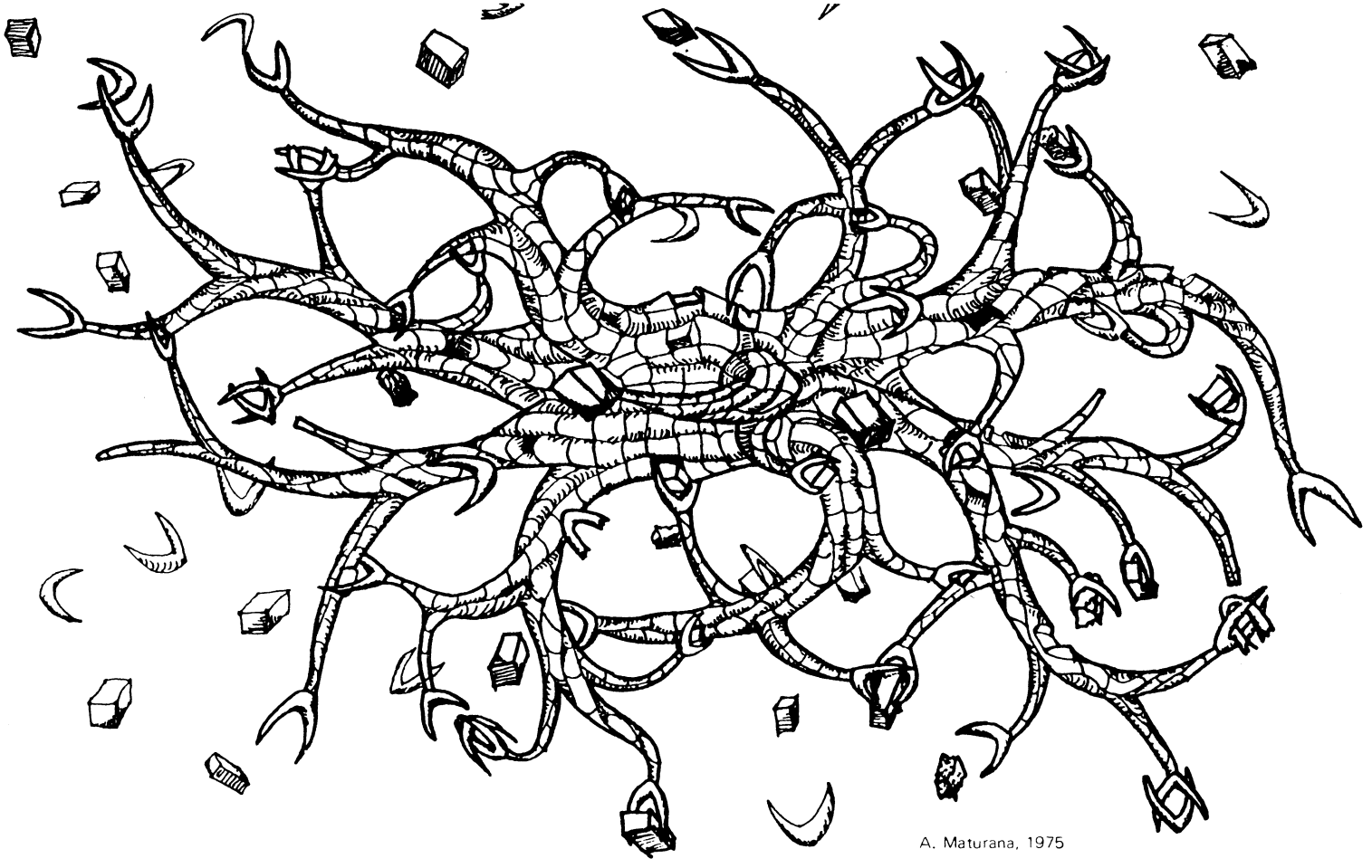
„Das Nervensystem als Ganzes ist so organisiert (organisiert sich so), daß es eine stabile Realität er-rechnet."

Aus dem Vorherigen erscheint es klar, daß wir es bei den „stabilen Realitäten" wieder mit einem Eigenwertproblem zu tun haben, und ich könnte mir vorstellen, daß diese Beobachtung in der Psychiatrie von Nutzen sein könnte.

In diesen Bemerkungen mögen manche ihre existentialistische Grundlage entdeckt haben. Mit der doppelten Schließung der Signalkreise – oder der vollständigen Schließung des Kausalkreises – habe ich ja weiter nichts erreicht, als die Autonomie eines jeden Lebewesens aufs neue zu stipulieren: Die Ursachen meiner Handlungen liegen nicht woanders oder bei jemand anderem – das wäre Heteronomie: der andere ist verantwortlich –, sondern die Ursachen meiner Handlungen liegen in mir, ich bin mein eigener Regler. Frankl, Jaspers oder Buber würden das vielleicht so sagen: In jedem Augenblick kann ich entscheiden, wer ich bin.

Aber damit fällt auch die Verantwortung für wer ich bin und wie ich handle auf mich zurück; Autonomie bedeutet Verantwortung, Heteronomie aber Verantwortungslosigkeit.

Da sehen wir, daß die erkenntnistheoretischen Probleme der Ethik mit denen der Kybernetik in einem weiten Bereich zusammenfallen, und damit fällt uns Kybernetikern die Verantwortung zu, uns an der Lösung der großen sozialen und ethischen Probleme unserer Zeit zu beteiligen.



A. Maturana, 1975

Bemerkungen zu einer Epistemologie des Lebendigen *

I Das Problem

Im ersten Viertel dieses Jahrhunderts sahen sich die Physiker und Kosmologen gezwungen, die grundlegenden Begriffe zu revidieren, die für die Naturwissenschaften bestimmend gewesen waren. Im letzten Viertel dieses Jahrhunderts dagegen werden die Biologen eine Revision all der Grundbegriffe erzwingen, die für die Wissenschaft schlechthin bestimmend sind. Nach jener „ersten Revolution" war klar, daß die klassische Vorstellung einer „letztgültigen Wissenschaft", d.h. einer objektiven Beschreibung der Welt, in der es keine Subjekte gibt (also eines „subjektlosen Universums"), in sich widersprüchlich ist.

Um die Widersprüche zu beseitigen, galt es, Position und Funktion des „Beobachters" (d.h. zumindest *eines* Subjekts) zu explizieren:

1. Beobachtungen sind nicht absolut, sondern relativ zum Standpunkt eines Beobachters (d.h. relativ zu seinem Koordinatensystem: Einstein);
2. Beobachtungen beeinflussen das Beobachtete und machen so jede Hoffnung des Beobachters zunichte, Vorhersagen treffen zu können (d.h. seine Unsicherheit ist absolut: Heisenberg).

Wir verfügen daher jetzt über die Binsenwahrheit, daß eine Beschreibung (des Universums) jemanden voraussetzt, der (es) beschreibt (beobachtet). Was wir nunmehr benötigen, ist die Beschreibung des „Beschreibers", oder mit anderen Worten, eine Theorie des Beobachters. Da gemäß dem Stande unseres Wissens nur lebende Organismen zu Beobachtern qualifiziert erscheinen, wird diese Aufgabe wohl den Biologen zufallen müssen. Aber auch der Biologe ist ein lebendes Wesen, und das bedeutet, daß er mit seiner Theorie nicht nur sich selbst, sondern auch das Schreiben dieser seiner Theorie erklären (können) muß. Damit ändern sich Art und Geltung wissenschaftlicher Aussagen, denn in traditioneller Auffassung wurde der Beobachter von seinen Beobachtungen getrennt, und alle Rückbezüglichkeit in wissenschaftlichen Aussagen mußte sorgfältig vermieden werden. Diese Trennung war keineswegs nur eine Marotte oder bloße Torheit, denn die Einbeziehung des Beobachters in seine Beschreibungen konnte unter bestimmten Umständen zu Paradoxa führen, wie die Äußerung „Ich bin ein Lügner" bestätigt.

Inzwischen ist jedoch mehr als klar geworden, daß diese enge Beschränkung nicht nur die mit wissenschaftlicher Tätigkeit verbundenen ethischen Probleme erzeugt, sondern auch die Erforschung des Lebens in seiner ganzen Fülle von molekularen bis hin zu sozialen Organisationen verkümmern läßt. Das Leben läßt sich nicht *in vitro* studieren, man muß es *in vivo* ergründen.

---

* Überarbeitete Fassung eines Vortrags vom 7. September 1972 im *Centre Royaumont pour une Science de l'Homme*, Royaumont, Frankreich. Den Rahmen bildete das internationale Kolloqium zum Thema "L'unité de l'homme: invariants biologiques et universaux culturels".

Die uns gestellte Frage „Die Einheit des Menschen: biologische Invarianten und kulturelle Universalien" kann nicht in der herkömmlichen beschränkten Weise angegangen werden, wenn unsere Antworten darauf dem heute erreichten Wissen um unsere eigene Biologie und Kultur entsprechen sollen.

Im Gegensatz zur klassischen Problemstellung wissenschaftlicher Forschung, die zunächst eine beschreibungsinvariante „objektive Welt" postuliert (als ob es so etwas gäbe) und sodann versucht, deren Beschreibung anzufertigen, sehen wir uns heute herausgefordert, eine beschreibungsinvariante „subjektive Welt" zu entwickeln, d. h. eine Welt, die den Beobachter einschließt. *Das ist das Problem.*

Durchaus in Übereinstimmung aber mit der klassischen Tradition wissenschaftlicher Forschung, die ständig „Wie?" fragt und nicht „Was?", verlangt diese Aufgabe eine Epistemologie des „Wie erkennen wir?" statt des „Was erkennen wir?"

Die folgenden Notizen zu einer Epistemologie für Lebewesen beschäftigen sich mit diesem „Wie?". Vielleicht können sie als Vergrößerungsglas dienen, um das Problem deutlicher erkennbar zu machen.

II Einführung

Die mit 1,2,3,...,12 numerierten zwölf Aussagen unter den folgenden 80 Notizen bilden den Minimalrahmen für das Bezugssystem, in dem die verschiedenen Konzepte, die erörtert werden, ihre Bedeutung gewinnen sollen. Da die Aussage 12 wieder direkt zur Aussage 1 zurückführt, können die Notizen im Kreis gelesen werden. Kommentare, Beweise und Erklärungen aber, die sich auf diese 12 Aussagen beziehen, sind diesen unter entsprechenden Dezimalzahlen (z.B. „5.423") angefügt; dabei bezieht sich die letzte Ziffer („3") auf eine Aussage, die durch die Ziffern davor bezeichnet ist („5.42") usw. („5.42" bezieht sich auf „5.4" usw.).

Obwohl man in diese Notizen an jeder beliebigen Stelle eintreten und sodann den ganzen Kreis durchlaufen kann, schien es ratsam, den Kreis zwischen den Aussagen „11" und „1" zu unterbrechen, die Notizen mit Aussage 1 zu beginnen und in entsprechender linearer Abfolge zu präsentieren.

Da der verwendete Formalismus manchem Leser eher verwirrend denn klärend erscheinen mag , schicke ich eine übersichtsweise Erläuterung der – etwas modifizierten – 12 Aussagen in Prosa voraus, um die Lektüre der Notizen selbst zu erleichtern.

1. *Die Umwelt wird erfahren als der Ort von Objekten, die stationär sind, die sich bewegen, oder die sich verändern.* ☚ [1])

Diese Aussage mag auf den ersten Blick harmlos erscheinen, beim zweiten Hinsehen aber wird man sich vielleicht nach der Bedeutung eines „sich verändernden Objekts" zu fragen beginnen. Meinen wir damit die Veränderung der Erscheinung(sweise) desselben Objekts, wie etwa dann, wenn wir einen Würfel drehen oder wenn sich ein Mensch umwendet, und wir beide dennoch weiter als dieselben Objekte (d.h. denselben Würfel, dieselbe Person usw.) ansehen? Oder wenn wir einen Baum wachsen sehen, oder einen alten Schulkameraden nach 10 oder 20 Jahren wieder treffen, und uns fragen, ob sie sich verändert haben oder dieselben

[1]) ☚ zeigt in dieser Einführung das Ende einer numerierten Aussage an.

geblieben sind, oder ob sie vielleicht nur in der einen Hinsicht anders, in der anderen aber unverändert sind? Oder wenn Circe Männer in Tiere verwandelt, oder ein Freund einen schlimmen Herzanfall erleidet, und wir uns fragen, was in diesen Metamorphosen invariant ist und was sich verändert? Wer sagt uns denn, daß es sich um dieselben Personen oder Objekte handelt?

Wir wissen aus Untersuchungen von Piaget (1954) und anderen (Witz/Easley 1972), daß „Objektkonstanz" eine von den vielen kognitiven Fertigkeiten ist, die in der frühen Kindheit erworben werden und folglich dem Einfluß von Sprache und Kultur unterliegen.

Um daher Begriffe wie „biologische Invarianten", „kulturelle Universalien" usw. sinnvoll verwenden zu können, müssen zunächst die logischen Eigenschaften von Begriffen wie „Invarianz" und „Veränderung" festgestellt werden.

Aus den folgenden Notizen wird klar werden, daß diese logischen Eigenschaften die Eigenschaften von Beschreibungen (Repräsentationen) sind und nicht die von Objekten. „Objekte", so zeigt sich, verdanken ihre Existenz in der Tat den Eigenschaften von Repräsentationen.

Hierzu werden die nächsten vier Aussagen formuliert.

2. *Die logischen Eigenschaften von „Invarianz" und „Veränderung" sind die Eigenschaften von Repräsentationen. Wird dies mißachtet, entstehen Paradoxa.*

Zwei Paradoxa, die sich ergeben, wenn die Begriffe „Invarianz" und „Veränderung" in einem kontextuellen Vakuum definiert werden, werden zitiert. Sie belegen die Notwendigkeit, Repräsentationen zu formalisieren.

3. *Die Repräsentationen* R *bzw.* S *werden mit Bezug auf zwei Mengen von Variablen* {x} *und* {t} *formalisiert, die versuchsweise als „Entitäten" bzw. „Augenblicke" benannt werden.*

Die Schwierigkeit anzufangen, über etwas zu reden, was erst später so verständlich wird, daß man beginnen kann, darüber zu reden, wird hier dadurch umgangen, daß zwei Mengen von noch undefinierten Variablen „versuchsweise" höchst bedeutsame Namen gegeben werden, nämlich „Entitäten" und „Augenblicke", die erst später begründet werden.

Diese scheinbare Abweichung vom exakten Vorgehen ist als Zugeständnis an die Verständlichkeit anzusehen. Die Streichung der bedeutungsvollen Etiketten dieser Variablen verändert den Argumentationsgang nicht.

Unter dieser Aussage werden Ausdrücke für Repräsentationen entwickelt, die verglichen werden können. Dies umgeht die scheinbare Schwierigkeit, einen Apfel mit sich selbst zu vergleichen, bevor und nachdem er geschält worden ist. Es bereitet aber nur geringe Schwierigkeiten, den geschälten Apfel, wie er *jetzt gesehen* wird, mit dem ungeschälten Apfel zu vergleichen, wie er in seinem früheren Zustand *erinnert* wird.

Mit dem Begriff „Vergleich" wird jedoch eine Operation („Rechnen") über Repräsentationen eingeführt, die eine genauere Analyse erfordert. Dies geschieht durch die nächste Aussage. Von da an wird der Begriff „Rechnen" konsistent auf alle (nicht notwendig numerischen) Operationen angewendet, die entweder Symbole (im „abstrakten" Sinne) oder deren physikalische Erscheinungsformen (im „konkreten" Sinne) transformieren, modifizieren, ordnen, neu anordnen usw. Dadurch soll das Gefühl dafür verstärkt werden, daß diese Operationen in der struk-

turellen und funktionalen Organisation vollentwickelter Nervengewebe oder anderswie konstruierter Maschinen verwirklicht werden können.

4. *Wir betrachten Relationen,* „Rel", *zwischen den Repräsentationen* R *und* S.

Es wird aber gleich eine höchst spezifische Relation untersucht, nämlich eine „Äquivalenzrelation" zwischen zwei Repräsentationen. Aufgrund der strukturellen Merkmale von Repräsentationen sind die zur Bestätigung oder Widerlegung der Äquivalenz von Repräsentationen nötigen Rechenvorgänge nicht trivial. Wenn man die Rechengänge für die Herstellung von Äquivalenzen verfolgt, entstehen „Objekte" und „Ereignisse als *Resultate* bestimmter Rechenprozesse, die als Prozesse der Abstraktion und des Merkens (Gedächtnisses) identifiziert werden.

5. *Objekte und Ereignisse sind keine primitiven Erfahrungen. Objekte und Ereignisse sind Repräsentationen von Relationen.*

Da „Objekte" und „Ereignisse" keine primären Erfahrungen sind und folglich keinen absoluten (objektiven) Status beanspruchen können, sind die Relationen zwischen ihnen, d.h. unsere „Umwelt", eine rein persönliche Angelegenheit, deren spezifische einschränkende Bedingungen in anatomischen und kulturellen Faktoren liegen. Damit löst sich das Postulat einer „externen (objektiven) Realität" auf, es entsteht eine Realität, die durch interne Rechenverfahren festgelegt wird (Castañeda 1971).

6. *Von einem operationalen Gesichtspunkt ist die Errechnung einer bestimmten Relation eine Repräsentation dieser Relation.*

Zwei Schritte von grundlegender Wichtigkeit für die in diesen Notizen vorgelegte Argumentation werden hier gleichzeitig getan. Der eine besteht darin, eine (Er-)Rechnung als eine Repräsentation anzusehen, der andere darin, erstmals „Rekursionen" einzuführen. Mit Rekursion ist gemeint, daß eine Funktion zu ihrem eigenen Argument gemacht wird. In der obigen Aussage 6 geschieht dies dadurch, daß die Errechnung einer Relation zwischen *Repräsentationen* erneut als eine Repräsentation betrachtet wird.

Obwohl nun die Ersetzung der Repräsentation einer Relation durch einen Rechenprozeß keine besonderen begrifflichen Schwierigkeiten erzeugen mag (die Lochkarte eines Computerprogramms, das die Rechenvorgänge zur Ermittlung einer Relation steuert, kann als passende Metapher dienen), scheint doch die Einführung rekursiver Ausdrücke logischen Schandtaten aller Art Tür und Tor zu öffnen.

Es gibt jedoch Mittel, nicht in solche Fallen zu tappen. Eines besteht darin, eine Notation zu entwickeln, die die Ordnung der Repräsentationen festhält. So kann z.B. „die Repräsentation einer Repräsentation einer Repräsentation" als eine Repräsentation dritter Ordnung $R^{(3)}$ betrachtet werden. Das gleiche gilt für Relationen höherer Ordnung n: $Rel^{(n)}$.

Nachdem nun die Begriffe der Repräsentationen bzw. Relationen höherer Ordnung eingeführt worden sind, werden ihre physikalischen Erscheinungsformen definiert. Da Repräsentationen und Relationen Rechenprozesse sind, sind ihre Erscheinungsformen „Spezialrechner", die „Repräsentoren" und „Relatoren" heißen. Die verschiedenen Ebenen des Rechnens werden dadurch auseinandergehalten, daß derartige Strukturen als Repräsentoren (Relatoren) n-ter Ordnung bezeichnet werden. Mit diesen Begriffen besteht nun die Möglichkeit, „Organismen" einzuführen.

7. *Ein lebender Organismus ist ein Relator dritter Ordnung, der die Relationen berechnet, die den Organismus als Ganzheit erhalten.*

Das ganze Arsenal rekursiver Ausdrücke wird nun auf die rekursive Definition lebender Organismen angewendet, wie sie zuerst von Maturana (1970a; 1970b) vorgeschlagen wurde und von Maturana und Varela in ihrem Modell der „Autopoiese" (1972) weiterentwickelt worden ist.

Mit Hilfe dieses Formalismus und der in früheren Aussagen entwickelten Begriffe ist es nunmehr möglich, eine Interaktion zwischen der internen Repräsentation eines Organismus von sich selbst mit der eines anderen Organismus zu erklären. Dies führt zu einer Theorie der Kommunikation, die auf einer rein konnotativen „Sprache" basiert. Die überraschende Eigenschaft einer derartigen Theorie wird in der folgenden Aussage 8 beschrieben.

8. *Ein Formalismus, der notwendig und hinreichend ist für eine Theorie der Kommunikation, darf keine primären Symbole enthalten, die Kommunikabilien repräsentieren (z. B. Symbole, Wörter, Botschaften usw.).*

Diese Aussage mag auf den ersten Blick völlig unhaltbar erscheinen, nach längerem Nachdenken aber wird sicherlich klar, daß eine Theorie der Kommunikation sich zirkulärer Definitionen schuldig macht, wenn sie Kommunikabilien voraussetzt, um Kommunikation abzuleiten.

Der Kalkül rekursiver Ausdrücke umgeht diese Schwierigkeit. Die Tragweite solcher Ausdrücke wird durch das (indefinit rekursive) personale Reflexivpronomen „Ich" exemplifiziert. Natürlich ist die magische Semantik derartiger infiniter Rekursionen seit geraumer Zeit bekannt, man denke nur an die Äußerung „Ich bin der ich bin" (Exodus 3, 14).

9. *Terminale Repräsentationen (Beschreibungen), die von einem Organismus hergestellt werden, manifestieren sich in seinen Bewegungen; die logische Struktur von Beschreibungen ergibt sich folglich aus der logischen Struktur von Bewegungen.*

Die zwei Grundmerkmale der logischen Struktur von Beschreibungen, nämlich ihr Sinn (Bejahung oder Verneinung) und ihr Wahrheitswert (Wahr oder Falsch), beruhen, so wird gezeigt, auf der logischen Struktur der Bewegung: Annäherung und Abwendung mit Bezug auf den ersten Aspekt, Funktionieren oder Versagen des bedingten Reflexes mit Bezug auf den zweiten.

Es ist nunmehr möglich, eine exakte Definition des Begriffs der „Information" einer Äußerung zu entwickeln. „Information" ist ein relatives Konzept, das Bedeutung nur dann annimmt, wenn es auf die kognitive Struktur des Beobachters dieser Äußerung (des „Rezipienten") bezogen wird.

10. *Die Information einer Beschreibung hängt von der Fähigkeit eines Beobachters ab, aus dieser Beschreibung Schlußfolgerungen abzuleiten.*

Die klassische Logik unterscheidet zwischen zwei Formen des Schließens: Deduktion und Induktion (Aristoteles, *Metaphysik*). Während es im Prinzip möglich ist, unfehlbare deduktive Schlüsse zu ziehen („Notwendigkeit"), ist es im Prinzip unmöglich, unfehlbare induktive Schlüsse zu ziehen („Zufall"). Zufall und Notwendigkeit sind folglich Begriffe, die sich nicht auf die Welt beziehen, sondern auf unsere Versuche, diese Welt (bzw. eine Beschreibung davon) zu erzeugen.

11. *Die Umwelt enthält keine Information: die Umwelt ist wie sie ist.*

12. *Zurück zu Aussage* 1.

III Bemerkungen

1. *Die Umwelt wird erfahren als der Ort von Objekten, die stationär sind, die sich bewegen, oder die sich verändern.*

1.1 „Veränderung" setzt Invarianz voraus, „Invarianz" Veränderung.

2. *Die logischen Eigenschaften von „Invarianz" und „Veränderung" sind die Eigenschaften von Repräsentationen. Wird dies mißachtet, entstehen Paradoxa.*

2.1 Das Paradoxon der „Invarianz":

WAS VERSCHIEDEN IST, SOLL DASSELBE SEIN

Aber es hat keinen Sinn zu schreiben: $x_1 = x_2$ (warum die Subskripte?). Und $x = x$ sagt zwar etwas über „=", nichts aber über x.

2.2 Das Paradoxon der „Veränderung":

WAS DASSELBE IST, SOLL VERSCHIEDEN SEIN

Aber es hat keinen Sinn zu schreiben: $x \neq x$.

3. *Wir formalisieren die Repräsentationen* R, S, ... *mit Bezug auf zwei Mengen von Variablen* x_i *und* t_j (, $j = 1, 2, 3, \ldots$), *die versuchsweise „Enitäten" bzw. „Augenblicke" heißen sollen.*

3.1 Die Repräsentation R einer Entität x mit Bezug auf den Augenblick t_1 ist verschieden von der Repräsentation dieser Entität mit Bezug auf den Augenblick t_2:

$$R(x(t_1)) \neq R(x(t_2)) .$$

3.2 Die Repräsentation S eines Augenblicks mit Bezug auf die Entität x_1 ist verschieden von der Repräsentation dieses Augenblicks mit Bezug auf die Entität x_2:

$$S(t(x_1)) \neq S(t(x_2)) .$$

3.3 Die vergleichende Beurteilung („verschieden von") kann jedoch nicht ohne einen Mechanismus vorgenommen werden, der diese Unterscheidungen errechnet.

3.4 Wir verkürzen die Notation durch

$$R(x_i(t_j)) \rightarrow R_{ij}$$
$$S(t_k(x_l)) \rightarrow S_{kl}$$

$(i, j, k, l = 1, 2, 3, \ldots)$.

4. *Wir betrachten die Relationen* Rel_μ *zwischen den Repräsentationen* R *und* S:

$$Rel_\mu (R_{ij}, S_{kl})$$

$(\mu = 1,2,3,\ldots)$.

4.1 Wir nennen die Relation, die die Unterscheidung $x_i \neq x_l$ und $t_j \neq t_k$ auslöscht (d. h. $i = l$; $j = k$), die „Äquivalenzrelation"; sie sei repräsentiert durch

$$Equ(R_{ij}, S_{ji}) .$$

4.1.1 Diese Formel repräsentiert eine Relation zwischen zwei Repräsentationen und lautet:

„Die Repräsentation R einer Entität x_i mit Bezug auf den Augenblick t_j ist äquivalent der Repräsentation S eines Augenblicks t_j mit Bezug auf die Entität x_i."

4.1.2 Eine mögliche linguistische Metapher für die oben angeführte Repräsentation der Äquivalenzrelation zwischen zwei Repräsentationen ist die Äquivalenzrelation zwischen einem „handelnden Ding" (in den meisten indo-europäischen Sprachen) und der „dingenden Handlung" (in einigen afrikanischen Sprachen) (kognitive Dualität). Beispiel:

„Das Pferd galoppiert" ⇌ „Der Galopp pferdet".

4.2 Die Errechnung der Äquivalenzrelation 4.1 hat zwei Verzweigungen:

4.2.1 Ein Zweig errechnet Äquivalenzen nur für x:

$$Equ(R_{ij}, S_{ki}) = Obj(x_i) .$$

4.2.1.1 Die Berechnungen entlang dieses Zweiges der Äquivalenzrelation werden „Abstraktionen" genannt: Abs.

4.2.1.2 Die Ergebnisse dieses Zweiges des Rechenprozesses werden gewöhnlich „Objekte" (Entitäten) genannt, und ihre Invarianz unter verschiedenen Transformationen (t_j, t_k,...) wird dadurch angezeigt, daß jedes Objekt ein verschiedenes, aber invariantes Etikett N_i (einen „Namen") erhält:

$$Obj(x_i) \rightarrow N_i .$$

4.2.2 Der andere Zweig errechnet Äquivalenzen ausschließlich für t:

$$Equ(R_{ij}, S_{jl}) \equiv Eve(t_j) .$$

4.2.2.1 Die Berechnungen entlang dieses Zweiges der Äquivalenzrelation werden „Merken" („Gedächtnis") genannt: Mem.

4.2.2.2 Die Ergebnisse dieses Zweiges des Rechenprozesses werden gewöhnlich als „Ereignisse" (Augenblicke) bezeichnet, und ihre Invarianz unter verschiedenen Transformationen (x_i, x_l,...) wird dadurch angezeigt, daß jedes Ereignis mit einem verschiedenen, aber invarianten Etikett T_j („Zeit") verknüpft wird:

$$Eve(t_j) \rightarrow T_j .$$

4.3 Damit ist gezeigt, daß die Begriffe „Objekt", „Ereignis", „Name", „Zeit", „Abstraktion", „Merken" („Gedächtnis"), „Invarianz", „Veränderung", einander erzeugen.

Daraus ergibt sich die nächste Aussage.

5. *Objekte und Ereignisse sind keine primitiven Erfahrungen. „Objekte" und „Ereignisse" sind Repräsentationen von Relationen.*

5.1 Eine mögliche graphische Metapher für die Komplementarität von „Ereignis" und „Objekt" ist ein rechtwinkeliges Gitter, das von beiden gebildet wird (Bild 1).

5.2 „Umwelt" ist die Repräsentation von Relationen zwischen „Objekten" und „Ereignissen":

$$Env(Obj, Eve) .$$

5.3 Da die Errechnung von Äquivalenzrelationen nicht einmalig ist, sind auch die Ergebnisse dieser Errechnungen, nämlich „Objekte" und „Ereignisse", nicht einmalig.

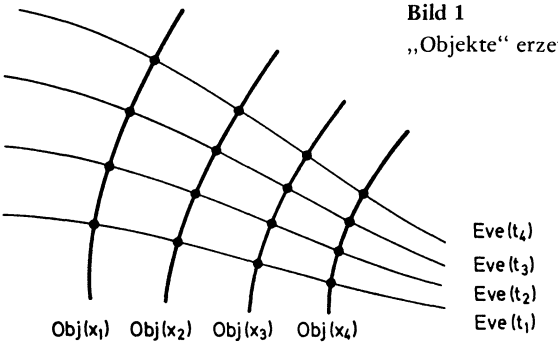


Bild 1
„Objekte" erzeugen „Ereignisse" und umgekehrt.

5.3.1 Dies erklärt die Möglichkeit einer beliebigen Anzahl verschiedener, intern jedoch konsistenter (sprachbedingter) Taxonomien.

5.3.2 Dies erklärt die Möglichkeit einer beliebigen Anzahl verschiedener, intern jedoch konsistenter (kulturbedingter) Wirklichkeiten.

5.4 Da die Errechnung von Äquivalenzrelationen über primitiven Erfahrungen ausgeführt wird, ist eine Außenwelt keine notwendige Voraussetzung für die Errechnung einer Wirklichkeit.

6. *Von einem operationalen Gesichtspunkt ist die Errechnung Cmp(Rel) einer bestimmten Relation eine Repräsentation dieser Relation:*

$R = Cmp(Rel)$.

6.1 Eine mögliche mathematische Metapher für die Äquivalenz einer Berechnung und einer Repräsentation ist etwa der Algorithmus von Wallis für das infinite Produkt:

$$2 \cdot \frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdot \frac{6}{5} \cdot \frac{6}{7} \cdots$$

Da dies eine der vielen möglichen Definitionen der Zahl π (3,14159...) ist, und da π eine Zahl ist, können wir π als eine (numerische) Repräsentation dieser Berechnung ansehen.

6.2 Wir wollen Repräsentationen von Berechnungen von Relationen „Repräsentationen zweiter Ordnung" nennen. Dies wird verständlich, wenn eine derartige Repräsentation voll ausgeschrieben wird:

$$R = Cmp(Rel(R_{ij}, S_{kl}))\ ,$$

wobei R_{ij} und S_{kl} natürlich, wie früher (3.3), „Repräsentationen erster Ordnung" sind.

6.2.1 Aus dieser Notation wird klar, daß Repräsentationen erster Ordnung als Relationen nullter Ordnung interpretiert werden können (man beachte die doppelten Subskripte bei S und R).

6.2.2 Aus dieser Notation wird außerdem klar, daß Repräsentationen und Relationen höherer Ordnung (n-ter Ordnung) formuliert werden können.

6.3 Ein physikalischer Mechanismus, der eine Repräsentation n-ter Ordnung (oder eine Relation n-ter Ordnung) errechnet, heiße ein „Repräsentor n-ter Ordnung", $RP^{(n)}$, bzw. ein „Relator n-ter Ordnung", $RL^{(n)}$).

6.4 Die externalisierte physikalische Manifestation des Ergebnisses einer Errechnung heiße eine „terminale Repräsentation" oder eine „Beschreibung".

6.5 Eine mögliche mechanische Metapher für Relator, Relation, Objekte und Beschreibungen ist ein mechanischer Tischrechner (der Relator), dessen interne Struktur (die Anordnung von Rädern und Stiften) die Repräsentation einer Relation ist, die gewöhnlich „Addition" genannt wird: Add(a,b). Gegeben zwei Objekte, a = 5, b = 7, berechnet diese Maschine eine terminale Repräsentation (eine Beschreibung) der Relation zwischen diesen zwei Objekten in digitaler dekadischer Form:

Add (5, 7; 12).

6.5.1 Natürlich kann eine Maschine mit einer anderen internen Repräsentation (Struktur) der gleichen Relation add(a, b; c) eine andere terminale Repräsentation (Beschreibung) dieser Relation zwischen denselben Objekten erzeugen, etwa in der Form des Produkts von Primzahlen:

Add $(5, 7; 2^2 \cdot 3^1)$.

6.6 Eine weitere mögliche mechanische Metapher für die Auffassung der Berechnung einer Relation als einer Repräsentation dieser Relation ist ein Elektronenrechner mit seinem Programm. Das Programm steht für die besondere Relation und verknüpft die Teile der Maschine auf solche Art, daß die terminale Repräsentation (der Ausdruck) des bearbeiteten Problems mit der gewünschten Form übereinstimmt.

6.6.1 Ein Programm, das Programme errechnet, wird als „Metaprogramm" bezeichnet. In unserer Terminologie ist eine Maschine, die Metaprogramme akzeptiert, ein Relator zweiter Ordnung.

6.7 Diese Metaphern bekräftigen die bereits früher gemachte Feststellung (5.3), daß die Berechnungen von Repräsentationen von Objekten und Ereignissen nicht einmalig sind.

6.8 Diese Metaphern machen darüber hinaus deutlich, daß mein Nervengewebe, das z.B. eine terminale Repräsentation in der Form der folgenden Äußerung errechnet: „Das ist die Brille meiner Großmutter", weder meiner Großmutter noch ihrer Brille ähnlich sieht, noch daß sich irgendeine „Spur" von beiden in ihm findet (ebensowenig wie es Spuren der Zahl „12" in den Rädern und Stiften eines Tischrechners oder Spuren irgendwelcher Zahlen in einem Computerprogramm gibt). Außerdem sollte meine Äußerung „Das ist die Brille meiner Großmutter" weder mit der Brille meiner Großmutter noch mit dem Programm verwechselt werden, das diese Äußerung errechnet, noch auch mit der Repräsentation (physikalischen Manifestation) dieses Programms.

6.8.1 Es läßt sich jedoch eine Relation zwischen der Äußerung, den Objekten und den diese beiden errechnenden Algorithmen berechnen (vgl. 9.4).

7. *Ein lebender Organismus* Ω *ist ein Relator dritter Ordnung* ($\Omega = RL^{(3)}$), *der die Relationen errechnet, die den Organismus als Ganzheit erhalten* (Maturana 1970a; 1970b):

$$\Omega\{Equ[R(\Omega(Obj)),\ S(Eve(\Omega))]\}.$$

Dieser Ausdruck ist rekursiv in Ω.

7.1 Ein Organismus ist für sich selbst das letztgültige Objekt.

7.2 Ein Organismus, der eine Repräsentation dieser Relation errechnen kann, hat Ich-Bewußtsein.

7.3 Unter den internen Repräsentationen der Errechnung von Objekten $Obj(x_i)$ innerhalb eines Organismus Ω kann es eine Repräsentation $Obj(\Omega^*)$ eines anderen Organismus Ω^* geben. Umgekehrt kann es in Ω^* eine Repräsentation $Obj^*(\Omega)$ geben, die Ω errechnet.

7.3.1 Beide Repräsentationen sind in Ω bzw. Ω^* rekursiv. Es gilt z.B. für Ω:

$$Obj^{(n)}(\Omega^{*(n-1)}(Obj^{*(n-1)}(\Omega^{(n-2)}(Obj^{(n-2)}(\ldots\Omega^*))))).$$

7.3.2 Dieser Ausdruck ist der Kern einer Theorie der Kommunikation.

8. *Ein Formalismus, der notwendig und hinreichend ist für eine Theorie der Kommunikation, darf keine primären Symbole enthalten, die Kommunikabilien repräsentieren (z. B. Symbole, Wörter, Botschaften usw.).*

8.1 Dies gilt, weil eine „Theorie" der Kommunikation, die primäre Kommunikabilien enthielte, keine Theorie, sondern eine Technologie der Kommunikation wäre, die Kommunikation als gegeben voraussetzt.

8.2 Die Aktivität des Nervensystems eines Organismus kann nicht mit einem anderen Organismus geteilt werden.

8.2.1 Dies bedeutet, daß in der Tat nichts „kommuniziert" wird bzw. werden kann.

8.3 Da der Ausdruck in 7.3.1 zyklisch werden kann (wenn $Obj^{(k)} = Obj^{(k-2i)}$), liegt es nahe, eine teleologische Theorie der Kommunikation zu entwickeln, in der das angestrebte Ziel darin besteht, $Obj(\Omega^*)$ unter allen Einwirkungen von seiten Ω^* invariant zu halten.

8.3.1 Es liegt auf der Hand, daß in einer solchen Theorie Fragen wie „Siehst Du die Farbe dieses Objekts so wie ich sie sehe?" irrelevant werden.

8.4 Kommunikation ist die Interpretation der Interaktion zwischen zwei Organismen Ω_1 und Ω_2 durch einen Beobachter.

8.4.1 Seien $Evs_1 \equiv Evs(\Omega_1)$ und $Evs_2 \equiv Evs(\Omega_2)$ *Sequenzen* von Ereignissen $Evs(t_j)$, $(j = 1,2,3,\ldots)$ mit Bezug auf zwei Organismen Ω_1 und Ω_2, und sei Com die (interne) Repräsentation einer Relation zwischen diesen Ereignissequenzen durch einen Beobachter:

$$OB(Com(Evs_1, Evs_2)).$$

8.4.2 Da Ω_1 oder Ω_2 oder beide Beobachter sein können ($\Omega_1 = OB_1$; $\Omega_2 = OB_2$), kann der obige Ausdruck in Ω_1 oder in Ω_2 oder in beiden rekursiv werden.

8.4.3 Dies zeigt, daß „Kommunikation" die (interne) Repräsentation einer Relation zwischen (einer internen Repräsentation von) mir selbst und jemand anders ist.

$$R(\Omega^{(n+1)},\ Com(\Omega^{(n)},\ \Omega^*)).$$

8.4.4 Wir kürzen dies ab als

$$C(\Omega^{(n)}, \Omega^*) .$$

8.4.5 In diesem Formalismus erscheint das personale Reflexivpronomen „Ich" als der (indefinit angewandte) rekursive Operator

$$\mathrm{Equ}[\Omega^{(n+1)}C(\Omega^{(n)}, \Omega^{(n)})]$$

oder in Worten:

„Ich bin die beobachtete Relation zwischen mir selbst und der Beobachtung meiner selbst."

8.4.6 „Ich" ist ein Relator (*und* Repräsentor) infiniter Ordnung.

9. *Terminale Repräsentationen (Beschreibungen), die von einem Organismus hergestellt werden, manifestieren sich in seinen Bewegungen; die logische Struktur von Beschreibungen ergibt sich folglich aus der logischen Struktur von Bewegungen.*

9.1 Es ist bekannt, daß die Anwesenheit eines wahrnehmbaren Agens in schwacher Konzentration einen Organismus veranlassen kann, sich auf dieses Agens hin zu bewegen (Annäherung). Das Auftreten des gleichen Agens in starker Konzentration kann diesen Organismus jedoch veranlassen, sich von ihm zu entfernen (Abwendung).

9.1.1 Dies bedeutet, daß „Annäherung" und „Abwendung" die Vorläufer von „Ja" und „Nein" sind.

9.1.2 Diese zwei Phasen elementaren Verhaltens, „Annäherung" und „Abwendung", bilden den operationalen Ursprung der beiden fundamentalen Axiome der zweiwertigen Logik, nämlich des „Gesetzes vom ausgeschlossenen Widerspruch":

$$\overline{x \;\&\; \bar{x}}$$

in Worten: „nicht: x *und* nicht-x",

und des „Gesetzes vom ausgeschlossenen Dritten":

$$x \vee \bar{x} ,$$

in Worten: „x *oder* nicht-x (vgl. Bild 2).

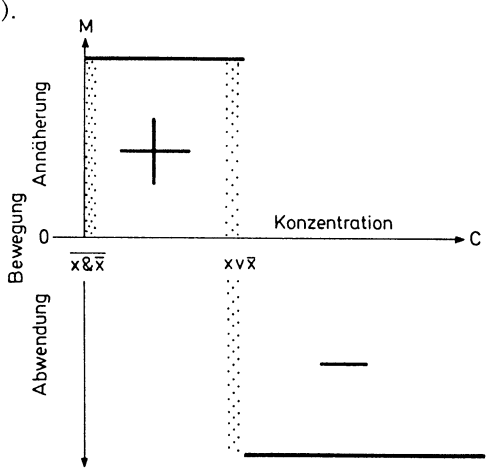


Bild 2
Die Gesetze vom „ausgeschlossenen Widerspruch" ($\overline{x \;\&\; \bar{x}}$) und vom „ausgeschlossenen Dritten" ($x \vee \bar{x}$) in den Übergangszonen zwischen Bewegungslosigkeit (M = 0) und Annäherung (+), und zwischen Annäherung (+) und Abwendung (−), als eine Funktion der Konzentration (C) eines wahrnehmbaren Agens.

9.2 Aus Wittgensteins *Tractatus* (1961) zitieren wir Aussage 4.0621:

„Daß aber die Zeichen ‚p' und ‚nicht-p' das gleiche sagen *können*, ist wichtig. Denn es zeigt, daß dem Zeichen ‚nicht' in der Wirklichkeit nichts entspricht.
Daß in einem Satz die Verneinung vorkommt, ist noch kein Merkmal seines Sinnes (nicht-nicht-p = p)."

9.2.1 Da in der Umwelt der Negation nichts entspricht, müssen die Negation ebenso wie alle anderen „logischen Partikel" (Inklusion, Alternation, Implikation usw.) innerhalb des Organismus selbst entstehen.

9.3 Beschreibungen können nicht nur im logischen Sinne bejahend oder verneinend, sondern auch wahr oder falsch sein.

9.3.1 Wir entnehmen dem Buch *Philosophy in a New Key* von Susanne K. Langer (1951):

„Der Gebrauch von Zeichen ist das allererste Zeugnis des Bewußtseins. Er ergibt sich ebenso früh in der biologischen Geschichte wie der berühmte ‚bedingte Reflex', bei dem ein begleitendes Merkmal eines Stimulus die Stimulusfunktion übernimmt. Das begleitende Merkmal wird zu einem *Zeichen* für die Situation, der die Reaktion genau angemessen ist. Dies ist der eigentliche Anfang des Denkens, denn hier liegt der Ursprung des *Irrtums* und somit der *Wahrheit*."

9.3.2 Damit ist nicht der Sinn (Ja oder Nein) von Beschreibungen, sondern auch deren Wahrheitswert (Wahr oder Falsch) an Bewegung (Verhalten) gekoppelt.

9.4 Sei D* eine terminale Repräsentation, hergestellt von einem Organismus Ω^*, und werde diese von einem Organismus Ω beobachtet; sei die interne Repräsentation der von Ω erzeugten Beschreibung $D(\Omega, D^*)$; und sei schließlich Ω's interne Repräsentation seiner Umwelt $E(\Omega, E)$. Dann ergibt sich:
Der Bereich der Relationen zwischen D und E, die von Ω errechnet werden können, repräsentiert die „Information", die von Ω gewonnen wird, wenn er Ω^* beobachtet:

$$\mathrm{Inf}(\Omega, D^*) \equiv \mathrm{Domain}\{\mathrm{Rel}_\mu(D, E)\}$$

($\mu = 1,2,3,\ldots,m$).

9.4.1 Der Logarithmus (zur Basis 2) der Anzahl m der Relationen Rel_μ, die von Ω errechenbar sind (oder der negative Mittelwert der logarithmischen Wahrscheinlichkeiten ihres Auftretens $\langle \log_2 p_i \rangle = \Sigma p_i \log_2 p_i$; $i = 1 \rightarrow m$), ist der „Informationsbetrag H" der Beschreibung D* mit Bezug auf Ω:

$$H(D^*, \Omega) = \log_2 m$$

$$(\text{oder: } H(D^*) = -\sum_1^m p_i \log_2 p_i).$$

9.4.2 Dies zeigt, daß Information ein relativer Begriff ist. Dasselbe gilt für H.

9.5 Einer Arbeit von Jerzy Konorski (1962) entnehmen wir:

„Es ist nicht so, wie wir aufgrund unserer Introspektion geneigt sind anzunehmen, daß der Empfang von Information und deren Nutzung zwei getrennte Prozesse sind, die auf beliebige Art miteinander kombiniert werden können; im Gegenteil, Information und ihre Nutzung sind untrennbar und bilden in Wirklichkeit einen einzigen Prozeß."

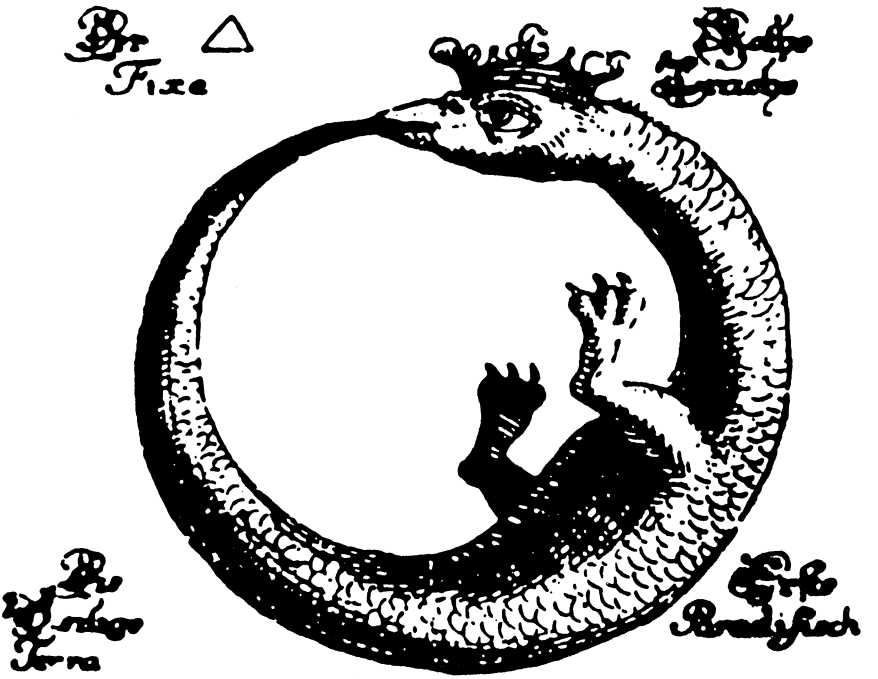
10. *Die Information einer Beschreibung hängt von der Fähigkeit eines Beobachters ab, aus dieser Beschreibung Schlußfolgerungen abzuleiten.*

10.1 „Notwendigkeit" entsteht aus der Fähigkeit, unfehlbare Deduktionen zu machen.

10.2 „Zufall" entsteht aus der Unfähigkeit, unfehlbare Induktionen zu machen.

11. *Die Umwelt enthält keine Information. Die Umwelt ist wie sie ist.*

12. *Die Umwelt wird erfahren als der Ort von Objekten, die stationär sind, die sich bewegen, oder die sich verändern (Aussage* 1).


Fixe

Gedanken und Bemerkungen über Kognition *

Gedanken

Es ist ein ganz übliches Verfahren, unser eigenes Bild in Dinge oder Funktionen von Dingen in der Außenwelt zu projizieren. Diese Art der Projektion nenne ich „Anthropomorphisierung". Da wir alle direkte Erfahrung von uns selbst haben, ist der unmittelbarste Weg zum Verstehen von X für uns, eine Abbildung zu finden, in der wir uns durch X selbst dargestellt sehen können. Dies wird deutlichst demonstriert dann, wenn wir bestimmten Dingen die Namen von Teilen unseres Körpers geben, weil sie diesen strukturell oder funktional ähnlich sind: Wir sprechen vom „Kopf" einer Schraube, von den „Backen" eines Schraubstocks, den „Zähnen" einer Säge, vom „Hals" einer Flasche, von den „Beinen" eines Tisches usw.

Die Surrealisten, die stets darauf aus waren, Ambivalenzen unserer kognitiven Prozesse deutlich zu machen, stellen uns solche drastisch vor Augen, indem sie sie semantisch korrekten Vorstellungen entgegensetzen: die Beine eines Stuhls (Bild 1, aus: Barr (ed.) 1947, 156), ein Brustkasten mit Schubladen (Bild 2, aus: Jean 1959, 284) usw.

Bild 1 „L'Ultra Meuble" (Kurt Seligman)

* Überarbeitete Fassung eines Forschungsberichts vom 2. März 1969 bei einem Symposium über "Cognitive Studies and Artificial Intelligence Research", finanziert von der Wenner-Gren Foundation for Anthropological Research, abgehalten am University of Chicago Center for Continuing Education, Chicago/Illinois.

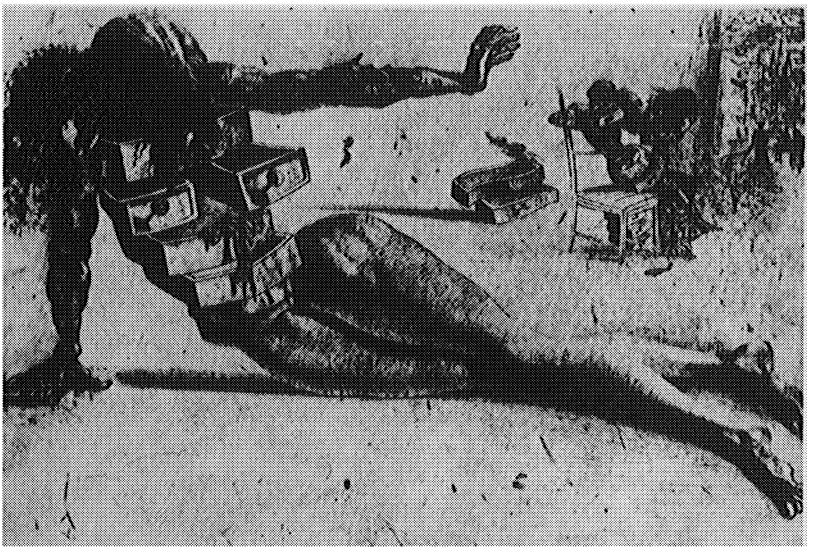


Bild 2 „Stadt der Schubladen" (Salvador Dali)

Um die Jahrhundertwende hatten die Tierpsychologen schwer gegen funktionalistische Anthropomorphismen in der Zoologie zu kämpfen, wo es von romantisierten Tieren mit menschlichen Eigenschaften nur so wimmelte. Da gab es den „treuen" Hund, das „edle" Pferd, den „stolzen" Löwen, den „schlauen" Fuchs usw. Konrad Lorenz, der große Ornithologe, wurde aus Wien hinausgejagt, als er unklugerweise vorschlug, die Population der sich allzu rasch vermehrenden unterernährten und tuberkuloseverseuchten Tauben der Stadt durch Falken zu kontrollieren, die die Eier der Tauben aus den Nestern holen sollten. Das goldene Wiener Herz konnte den Gedanken des „Taubenkindermordes" allerdings nicht ertragen. Die Wiener gaben den Tauben lieber doppelt soviel Futter. Als Lorenz darauf hinwies, daß sie dadurch nur doppelt so viele unterernährte und tuberkuloseverseuchte Tauben heranzüchteten, mußte er gehen, und zwar schnell!

Natürlich ist im Prinzip gegen Anthropomorphisierungen überhaupt nichts einzuwenden: In den meisten Fällen dienen sie als hilfreiche Algorithmen zur Steuerung des Verhaltens. Wenn man es mit einem Fuchs zu tun hat, ist es sicherlich vorteilhaft zu wissen, daß er „schlau" ist, d.h., daß er eine Herausforderung für unser Gehirn ist und nicht für unsere Muskeln.

Heute lebt der größte Teil der Menschheit in Städten und hat den direkten Kontakt mit der Tierwelt verloren. Stattdessen sind Möbelstücke aus Stahl mit gewissen funktionalen Merkmalen zu Gegenständen unserer Zuwendung geworden: die Computer. Und sie sind es nun, die mit romantisierenden Beiwörtern geschmückt werden. Da wir in einem Zeitalter der Naturwissenschaften und der Technik leben, nicht in einem des Gefühls und der Empfindsamkeit, sind die liebevollen Epitheta dieser unserer Maschinen nicht solche des Charakters, sondern solche des Intellekts. Auch wenn es durchaus möglich und vielleicht sogar angemessen ist, von einem „stolzen" IBM 360-50-System, von der „edlen" 1800

oder der „schlauen" PDP 8 zu sprechen, habe ich nie jemand diese Ausdrucksweise benutzen hören. Stattdessen romantisieren wir angebliche geistige Leistungen dieser Maschinen, wir sprechen von ihrem „Gedächtnis" und sagen, daß diese Maschinen „Information" speichern und wieder auffinden, daß sie „Probleme lösen", „Theoreme beweisen" usw. Man hat es hier augenscheinlich mit ziemlich intelligenten Geschöpfen zu tun, und es hat daher einige Versuche gegeben, einen A.I.Q, d.h. einen „artificial intelligence quotient", zu bestimmen, um auch in dieses neue Gebiet der „künstlichen Intelligenz" mit Kompetenz und Nachdruck irreführende Vorstellungen einzuführen, wie sie auch bei manchen prominenten Behavioristen noch gang und gäbe sind.

Unser rationales Verhältnis zu diesen Maschinen bedarf also einer klärenden Bestimmung, im emotionalen Bereich dagegen scheinen wir keine Probleme zu haben. Die folgenden Anmerkungen sollen lediglich die charmanten Ausführungen von Madeleine Mathiot (1970) ergänzen, in denen sie über verschiedene Grade der „Ehrfurcht und Bewunderung" berichtet, die mit den Personalpronomen „er", „sie" und „es" assoziiert werden. Sie entwickelt ein dreiwertiges logisches Stellenwertsystem, in dem das sächliche „es" keinen Bezug zu Ehrfurcht und Bewunderung zeigt, und zwar weder im negativen (Fehlen) noch im positiven Sinn (Vorhandensein), während das personale „er" bzw. „sie" durchaus entsprechend markiert ist. Das maskuline „er" ist durch das Fehlen von Ehrfurcht und Bewunderung gekennzeichnet, das feminine „sie" natürlich durch ihr Vorhandensein.

Als in den frühen 50er Jahren an der Universität von Illinois der Rechner ILLIAC II gebaut wurde, haben wir ihn alle mit dem Pronomen „es" bezeichnet. Die gegenwärtig an ILLIAC III arbeitende Computergruppe verspricht uns, daß „er" bald seine Arbeit aufnehmen wird. ILLIAC IV eröffnet aber bereits ganz andere Dimensionen: Die Planer meinen, daß das Rechnerpotential der Welt sich verdoppeln wird, sobald „sie" eingeschaltet werden kann.

Auch diese Anthropomorphismen sind selbstverständlich völlig in Ordnung, denn sie können die Entwicklung guter Arbeitsbeziehungen mit diesen Werkzeugen erleichtern. Da die meisten der Leute, die ich in unserer Computerabteilung kenne, männlich und heterosexuell sind, ist völlig klar, daß sie ihre Tage und Nächte bei der Arbeit lieber mit einer „Sie" als mit einem „Es" verbringen.

Im letzten Jahrzehnt ist jedoch etwas ganz Seltsames und Betrübliches geschehen: Nicht nur die Ingenieure, die mit solchen Systemen arbeiten, haben langsam zu glauben begonnen, daß die geistigen Funktionen, deren Namen zunächst *metaphorisch* gewissen Maschinenoperationen zugewiesen wurden, tatsächlich diesen Maschinen eigen sind, vielmehr haben auch gewisse Biologen – bewogen dadurch, daß es keine umfassende Theorie des Denkens gibt – zu glauben begonnen, daß bestimmte Maschinenoperationen, die unglücklicherweise die *Namen* einiger geistiger Prozesse tragen, diesen geistigen Operationen funktional tatsächlich isomorph sind. Auf der Suche nach der physiologischen Basis des Gedächtnisses begannen sie z.B. nach neuronalen Mechanismen zu suchen, die gewissen elektromagnetischen oder elektrodynamischen Mechanismen analog sein sollen, welche zeitabhängige Konfigurationen (Magnetbänder, Magnettrommeln oder Magnetkerne) oder räumliche Konfigurationen (Hologramme) des elektromagnetischen Feldes so „einfrieren", daß sie zu späterer Zeit erneut bearbeitet werden können.

Der Irrglauben an eine funktionale Isomorphie bei voneinander völlig unabhängigen und verschiedenartigen Prozessen, nur weil diese Prozesse mit dem gleichen *Namen* bezeichnet werden, ist in diesen beiden Fachdisziplinen heute so tief eingewurzelt, daß jeder, der nach dem Beispiel von Lorenz die Maschinen zu „de-anthropomorphisieren" und die Menschen zu „de-mechanisieren" sucht, ähnlich feindselige Einstellungen und Handlungen hervorruft, wie sie Lorenz entgegengebracht wurden, als er begann, die Tiere wieder zu Tieren zu machen, sie zu „animalisieren".

Andererseits ist der Widerstand gegen ein theoretisches Modell, in dem scheinbar voneinander trennbare höhere geistige Fähigkeiten, wie z.B. „Lernen", „Merken" („Gedächtnis"), „Wahrnehmen", „Vorhersagen" usw., als unterschiedliche Manifestationen eines einzigen umfassenden Phänomens „Kognition" aufgefaßt werden, durchaus verständlich. Dies würde nämlich bedeuten, daß man die bequeme Position aufzugeben hat, von der aus alle diese Fähigkeiten isoliert bearbeitet und auf ziemlich triviale Mechanismen reduziert werden können. Das Gedächtnis etwa wird, isoliert betrachtet, auf „Aufzeichnen" reduziert, „Lernen" auf „Veränderung", „Wahrnehmung" auf „Input" usw. Mit anderen Worten, durch die Abtrennung dieser Funktionen von der Totalität kognitiver Prozesse ist man das ursprüngliche Problem losgeworden und sucht nun nach Mechanismen, die ganz andere Aufgaben erfüllen. Diese können mit den Prozessen zusammenhängen oder auch nicht, die nach den Ausführungen Maturanas (1970a) der Erhaltung des Organismus als einer funktionierenden Einheit dienen.

Vielleicht können die folgenden drei Beispiele diese Problematik noch deutlicher machen.

Ich möchte mit dem „Gedächtnis" (engl.*memory*)* beginnen. Wenn Ingenieure über das „Gedächtnis" eines Computers sprechen, dann meinen sie nicht eigentlich das „Gedächtnis" eines Computers, sondern Vorrichtungen oder Systeme von Apparaturen, mit denen elektrische Signale festgehalten werden, so daß sie dann, wenn sie für weitere Manipulationen benötigt werden, erneut abgerufen werden können. Diese technischen Vorrichtungen sind daher Speicher oder Speichersysteme und zeigen die charakteristischen Merkmale aller Speicher, nämlich die Erhaltung der Qualität dessen, was in einem Zeitpunkt gespeichert und zu einem späteren Zeitpunkt wieder abgerufen wird. Der Inhalt dieser Speicher ist eine Aufzeichnung, englisch „record", und eben dies war in den Zeiten vor der Großen Semantischen Konfusion auch der englische Ausdruck für jene dünnen schwarzen Scheiben, die die Musik wiedergeben, welche auf ihnen aufgezeichnet ist. Ich stelle mir vor, welch große Augen der Verkäufer in einem Musikgeschäft machen würde, wenn man ihn nach dem „Gedächtnis" von Beethovens fünfter Symphonie fragte. Er würde den Kunden wahrscheinlich an den Buchladen nebenan verweisen. Und völlig zu recht, denn Aufzeichnungen vergangener Erfahrungen reproduzieren nicht die Ursachen dieser Erfahrungen, sie transformieren diese Erfahrungen vielmehr – durch einen Wechsel der qualitativen Bereiche – mit Hilfe einer Menge komplexer Prozesse in Äußerungen oder andere Formen symbolischen oder zielorientierten Verhaltens. Wenn man mich fragt, was ich zum Frühstück gegessen

* Anm. d. Übs.: Das englische Wort für den „Speicher" eines Computers, "memory", ist jedermann primär in seiner Bedeutung „Gedächtnis" vertraut, – daher die im folgenden dargelegten fatalen Konsequenzen.

habe, dann produziere ich nicht Rührei, sondern *sage* lediglich „Rührei". Das „Gedächtnis" eines Computers hat nichts mit solchen Transformationen zu tun, und es sollte auch nie damit zu tun haben. Das bedeutet jedoch nicht, daß ich meine, diese Maschinen würden eines Tages nicht ihre eigenen Memoiren schreiben. Um sie aber so weit zu bringen, müssen wir erst noch einige epistemologische Probleme lösen, denn erst dann können wir uns mit dem Problem der Konstruktion der dafür notwendigen Software und Hardware beschäftigen.

Auch das schmückende Beiwort „Problemlöser" – ebenso wie das Wort „Gedächtnis" für Aufzeichnungsapparate – ist eine irreführende Metapher für unsere Rechenmaschinen. Natürlich sind diese keine Problemlöser, denn sie haben überhaupt keine Probleme. Es sind *unsere* Probleme, die sie uns lösen helfen wie jedes andere Werkzeug auch, etwa ein Hammer, der in der Tat als „Problemlöser" dafür angesehen werden kann, Nägel in ein Brett zu schlagen. Die Gefahr einer solchen subtilen semantischen Verdrehung, durch die die Verantwortung für die eigenen Handlungen vom Menschen auf die Maschine geschoben wird, liegt darin, daß wir das Problem der Kognition aus dem Auge verlieren. Dadurch, daß wir der Verführung erliegen zu glauben, daß das Problem darin besteht, Lösungen für bestimmte wohldefinierte Fragestellungen zu finden, vergessen wir zunächst überhaupt zu fragen, was ein „Problem" ausmacht, sodann worin seine „Lösung" besteht, und schließlich – wenn ein Problem identifiziert ist –, warum wir es überhaupt lösen wollen.

Ein weiterer Fall pathologischer Semantik – und das letzte Beispiel meiner Polemik – ist der weitverbreitete Mißbrauch des Begriffs „Information". Dieses arme Ding wird heutzutage „verarbeitet", gespeichert", „wieder herbeigeschafft", „komprimiert", „zerlegt" usw., so als ob es Hackfleisch wäre. Da die Fallgeschichte dieser modernen Krankheit leicht einen ganzen Band füllen könnte, greife ich nur die sogenannten „Systeme der Speicherung und Wiederbereitstellung von Information" heraus, die etwa in der Form bestimmter fortgeschrittener bibliothekarischer Such- und Liefersysteme, computergestützter Datenverarbeitungssysteme, des Educational Resources Information Center (ERIC) usw. ganz ernsthaft als Analogmodelle für das Funktionieren des Gehirns vorgeschlagen worden sind.

Natürlich speichern diese Systeme keinerlei Information, sie speichern Bücher, Bänder, Mikrofiches oder andere Arten von Dokumenten, und es sind eben diese Bücher, Bänder, Mikrofiches oder anderen Dokumente, die wieder hervorgeholt werden und die nur dann die gewünschte Information liefern, wenn ein menschliches Bewußtsein sie entsprechend verarbeitet. Diese Sammlungen von Dokumenten „Systeme der Speicherung und Wiederbereitstellung von Information" zu nennen, ist ebenso falsch wie eine Garage als „System der Speicherung und Wiederbereitstellung von Transport" zu bezeichnen. Die Verwechslung von *Behältern* für potentielle Information mit der *Information* selbst führt wiederum dazu, das Problem der Kognition wunderschön in den blinden Fleck des wissenschaftlichen Sehfeldes zu rücken, so daß es, wie gewünscht, verschwindet. Wäre das Gehirn in der Tat mit irgendeinem dieser Systeme der Speicherung und Wiederbereitstellung vergleichbar und wäre es von diesen nur nach seiner Speicherkapazität und nicht auch nach der Qualität seiner Prozesse verschieden, dann würde eine entsprechende Theorie des Gehirns einen Dämon mit immensen kognitiven Fähigkeiten postulieren müssen, der durch dieses riesige System rast, um aus

dessen Beständen jeweils genau die Information zu extrahieren, die für den Besitzer des Gehirns gerade lebensnotwendig ist.

Difficile est satiram non scribere. Ganz offensichtlich bin ich mit dieser Schwierigkeit nicht fertiggeworden, und ich fürchte, daß es mir ebenso wenig gelingen wird, die andere Schwierigkeit zu bewältigen, nämlich nunmehr zu sagen, was Kognition *wirklich* ist. Es fällt mir sogar schwer, die Tiefe dieses unseres Problems angemessen in Worte zu fassen, wenn ich es in seinem vollen Umfang ernstnehme und vergegenwärtige. In einer Gruppe wie der unsrigen gibt es wahrscheinlich ebenso viele Arten, es zu betrachten, wie es Augenpaare gibt. Mich fasziniert jedenfalls nach wie vor das unergründliche und rätselhafte Geschehen, das sich ereignet, wenn Hans, Peters Freund, die Geräusche hört, die erzeugt werden, wenn jemand die folgenden schwarzen Zeichen liest:

ANNA IST PETERS SCHWESTER

– oder wenn Hans diese Zeichen auch nur sieht. Hans „weiß" danach, daß Anna Peters Schwester ist und verändert faktisch seine gesamte Einstellung gegenüber der Welt, wie sie sich aus seiner neuen Einsicht in eine relationale Struktur von Elementen dieser Welt ergibt.

Ich meine, daß wir die „kognitiven Prozesse", die diese Art von Wissen aus bestimmten Sinneswahrnehmungen erzeugen, noch keineswegs verstehen. Ich werde mir in diesem Augenblick nicht weiter den Kopf darüber zerbrechen, ob diese Sinneswahrnehmungen nun durch eine Interaktion des Organismus mit Objekten in der Welt oder mit deren symbolischen Repräsentationen verursacht werden. Wenn ich nämlich H. Maturana richtig verstehe, werden sich diese zwei Probleme bei angemessener Formulierung auf ein und dasselbe Problem reduzieren, nämlich auf das Problem der Kognition *per se*.

Um diese ganze Problematik für mich selber zu erklären, habe ich die folgenden Bemerkungen zusammengestellt, die um sechs Aussagen gruppiert sind, angeordnet nach $n = 1 \rightarrow 6$. Die mit n.1, n.2, n.3 etc. numerierten Aussagen sind Erläuterungen zu den Aussagen unter der Ziffer n. Die mit n.m1, n.m2 usw. bezeichneten Aussagen sind Kommentare zu den jeweiligen Aussagen n.m usw.

Hier sind sie.

Bemerkungen

1. Ein lebender Organismus, Ω, ist eine abgegrenzte, autonome Einheit, deren funktionale und strukturelle Organisation durch die Interaktion seiner miteinander verbundenen elementaren Konstituenten bestimmt wird.

1.1 Die elementaren Konstituenten, die Zellen, sind ihrerseits abgegrenzte, funktionale und strukturelle Einheiten, sie sind jedoch nicht notwendigerweise autonom.

1.1.1 Die Autonomie der Zellen geht mit der zunehmenden Differenzierung in Organismen von steigender Komplexität verloren, wodurch andererseits die angemessene „organische Umwelt" dafür geschaffen wird, daß diese komplexen Ganzheiten ihre strukturelle und funktionale Integrität erhalten können.

1.2 Ein lebender Organismus, Ω, wird durch eine geschlossene orientierbare Oberfläche begrenzt. Topologisch ist diese Oberfläche äquivalent einer Kugel mit einer geraden Anzahl von 2p Löchern, die paarweise durch Röhren verbunden sind. Die Zahl p wird als Genus der Oberfläche bezeichnet.

1.2.1 Wird die histologische Unterscheidung zwischen Ektoderm und Endoderm aufrechterhalten, dann ist eine Oberfläche vom Genus $p = (s + t)/2$ äquivalent einer Kugel mit s Oberflächenlöchern, die durch ein Netzwerk von Röhren mit t T-Verzweigungen verbunden sind. Das Ektoderm wird dann durch die Oberfläche der Kugel repräsentiert, das Endoderm durch die Innenwände der Röhren (Bild 3).

1.3 Jede geschlossene orientierbare Oberfläche ist metrisierbar. Daher kann jeder Punkt dieser Oberfläche durch die zwei Koordinaten α, β eines geodätischen Koordinatensystems bezeichnet werden, das die Oberfläche in geeigneter Weise bedeckt. Eine der Eigenschaften eines geodätischen Koordinatensystems liegt darin, daß es lokal cartesisch ist. Die Oberflächenkoordinaten α, β heißen „Eigen-Koordinaten" und werden durch ein einziges Symbol, ξ, bezeichnet.

1.3.1 Wenn für die Umgebung jedes Oberflächenpunktes ξ die Gaußsche Krümmung γ gegeben ist, dann legt die Gesamtheit der Tripel α, β, γ die Gestalt Γ der Oberfläche fest ($\gamma = \gamma[\alpha, \beta]$).

1.3.2 Da ein lebender Organismus durch eine geschlossene orientierbare Oberfläche begrenzt wird, läßt sich über die Oberfläche dieses Organismus in einem willkürlich gewählten „Ruhezustand" ein geeignetes geodätisches Koordinatensystem legen, und jedes Oberflächenelement (d.h. jede ektodermale oder endodermale Zelle) läßt sich durch die Eigen-Koordinaten ξ hinsichtlich seiner Lage kennzeichnen.

1.3.3 Eine entsprechend bestimmte Zelle c_ξ soll diesen ihren Namen nach aufeinanderfolgenden Verzerrungen der Oberfläche (kontinuierlichen Verzerrungen) und auch nach Transplantierung an den Ort ξ' (diskontinuierliche Verzerrungen) beibehalten.

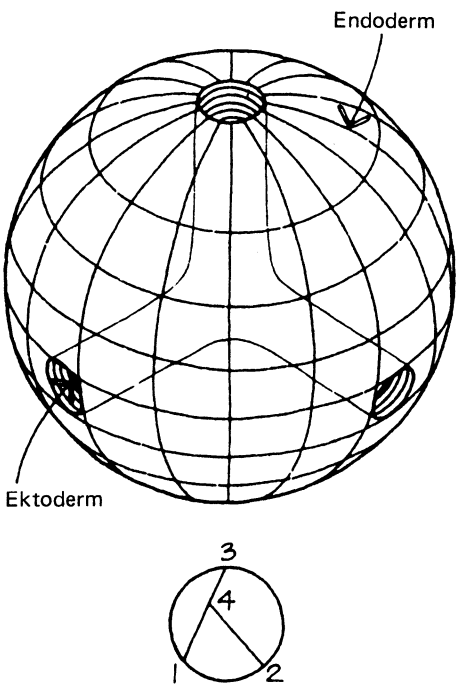


Bild 3
Geschlossene orientierbare Oberfläche vom Genus p = 2 (s = 3, t = 1, [s + t]/2 = 4/2 = 2)

1.3.3.1 Die geodätischen Koordinaten auf der Oberfläche des Organismus lassen sich auf eine topologisch äquivalente Einheitskugel (R = 1) abbilden, so daß jedem Punkt ξ und seiner Umgebung auf dem Organismus präzise ein Punkt λ mit seiner Umgebung auf der Einheitskugel entspricht. Jede Zelle c_ξ auf der Oberfläche des Organismus hat folglich eine Abbildung c_λ auf der Oberfläche der Einheitskugel.

1.3.3.2 Es ist klar, daß Oberflächenverzerrungen des Organismus, ja sogar Transplantierungen von Zellen von einem Ort zu einem anderen, sich nicht in irgendwelchen Veränderungen der Oberfläche dieser Kugel widerspiegeln. Die einmal hergestellte Abbildung bleibt gegenüber solchen Transformationen invariant, und diese Kugel wird daher im folgenden als „repräsentative Körperkugel" bezeichnet (Bild 3, oder geeignete Modifikationen mit $p > 2$).

1.3.4 Da das von einer geschlossenen orientierbaren Oberfläche umfaßte Volumen metrisierbar ist, gilt all das, was (1.3 → 1.3.3.2) über die Oberflächenpunkte ξ und die Zellen c_ξ gesagt worden ist, auch für die Rauminhaltspunkte ζ und die Zellen c_ζ bzw. die repräsentativen Zellen c_μ in der Körperkugel.

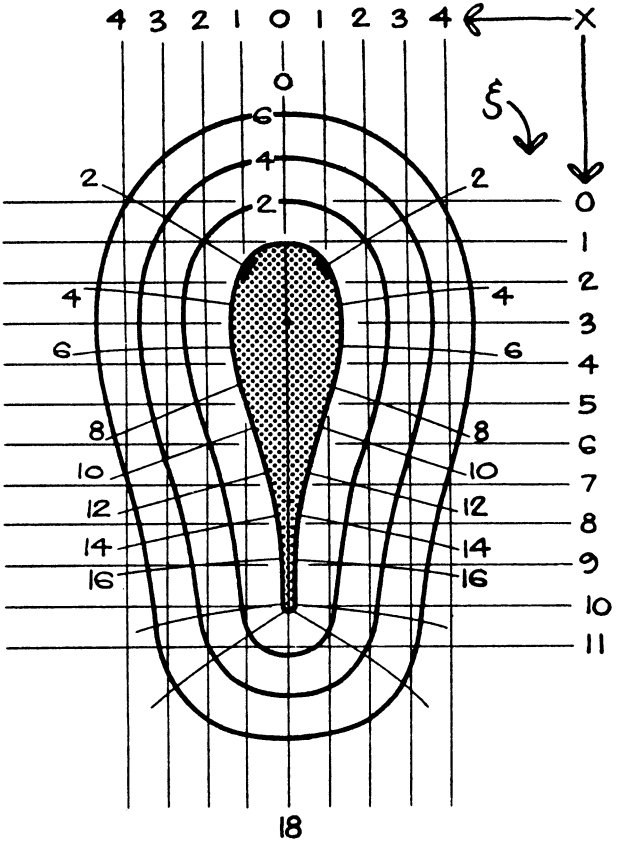


Bild 4a Geodäsie des Eigen-Koordinatensystems eines Organismus Ω im Ruhezustand [$\delta\Gamma = 0$], der in eine Umwelt mit euklidischer Metrik eingebettet ist.

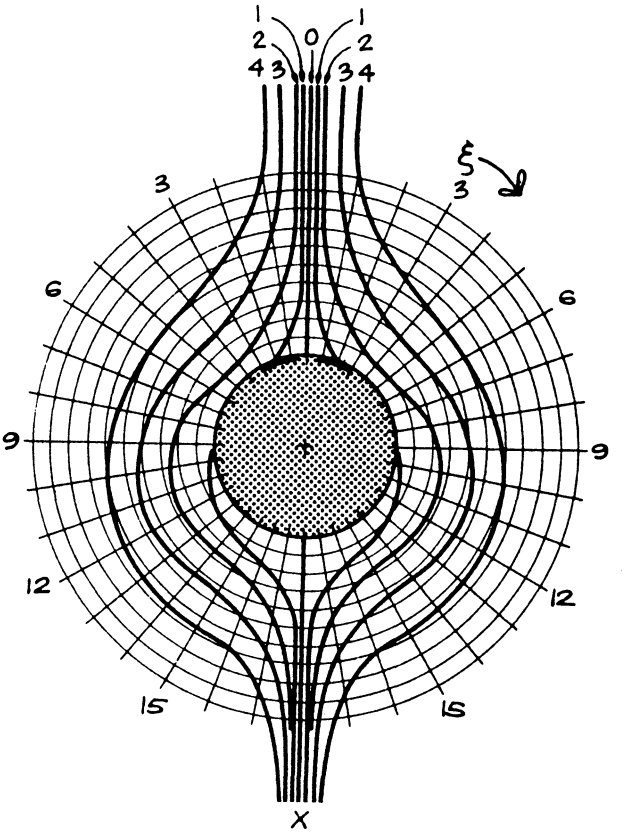


Bild 4b Geodäsie (Kreis, Radien) des Eigen-Koordinatensystems hinsichtlich der repräsentativen Körperkugel, die in eine Umwelt mit nicht-euklidscher Metrik eingebettet ist und dem Organismus im Ruhezustand entspricht (Bild 4a).

1.4 Der Organismus, Ω, wird in eine „Umwelt" mit festgelegter euklidischer Metrik eingebettet, d.h. mit den Koordinaten a,b,c, abgekürzt x, in der seine Position durch die Angabe dreier Umweltpunkte x_1, x_2 und x_3 mit drei Oberflächenpunkten ξ_1, ξ_2, ξ_3 des Organismus definiert wird. Umgekehrt wird die repräsentative Körperkugel in eine „repräsentative Umwelt" mit variabler nicht-euklidischer Metrik eingebettet, *mutatis mutandis* bzgl. der übrigen Bedingungen.

1.4.1 Die zwei Paare von Abbildungen (Bilder 4a und 4b sowie Bilder 5a und 5b) illustrieren die Konfiguration des Eigenraumes, ξ, eines Organismus (eines fischähnlichen Tieres), wie er von einer euklidischen Umwelt her (4a und 5a) gesehen wird, und die Konfiguration der nicht-euklidischen Umwelt, wie sie von der Einheitskugel (4b, 5b) für die zwei Fälle gesehen wird, in denen der Organismus in Ruhestellung (4a, 4b) bzgl. der übrigen Bedingungen.

2. Phylogenetisch und ontogenetisch entwickelt sich das Neuralrohr aus dem Ektoderm. Die Rezeptorzellen r_ξ sind differenzierte ektodermale Zellen c_ξ. Das gleiche gilt für die anderen Zellen tief im Körperinneren, die mitwirken an der

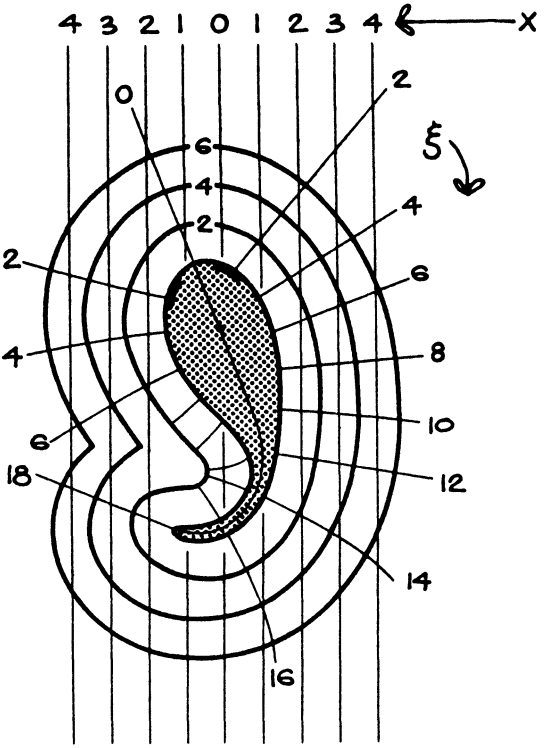


Bild 5a
Geodäsie des Eigen-Koordinatensystems eines Organismus Ω in Bewegung ($\delta\Gamma \neq 0$), der in eine Umwelt mit euklidischer Metrik eingebettet ist.

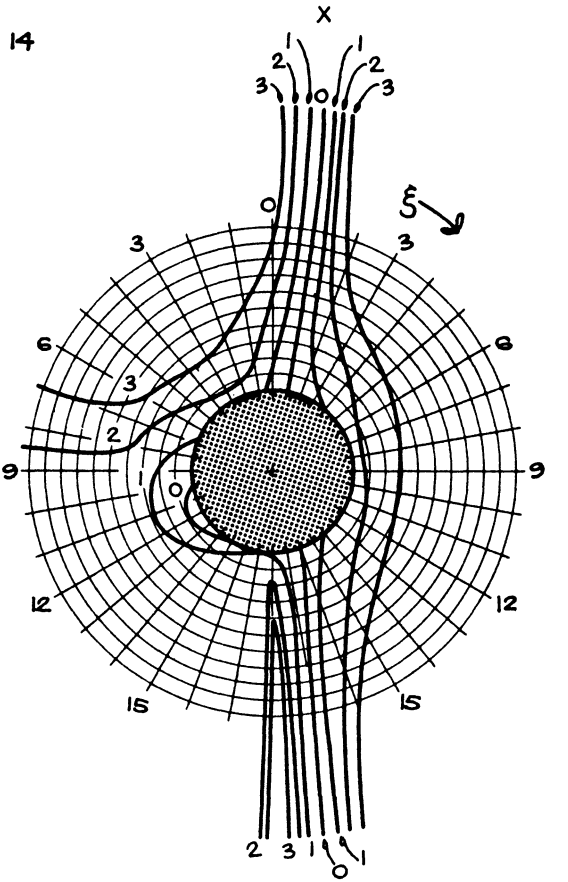


Bild 5b
Geodäsie (Kreise, Radien) des Eigen-Koordinatensystems hinsichtlich der repräsentativen Körperkugel, die in eine Umwelt mit nicht-euklidischer Metrik eingebettet ist und dem Organismus in Bewegung entspricht (Bild 5a).

Übertragung von Signalen (Neuronen) n_ξ und an der Erzeugung von Signalen (Propriozeptoren) p_ξ; es gilt schließlich für alle jene (Effektoren) e_ξ, die durch ihre Signale spezialisierte Fasern (Muskeln) m_ξ kontrahieren und somit Veränderungen $\delta\Gamma$ der Gestalt des Organismus verursachen (Bewegung).

2.1 Sei A ein Agens von der Menge A, das in der Umwelt verteilt ist und dessen Konzentration (Intensität) durch eine Verteilungsfunktion über einen Parameter p bestimmt ist:

$$S(x,p) = \frac{d^2A}{dxdp} = \left(\frac{da}{dp}\right)_x$$

wobei

$$\int_0^\infty S(x,p)dp = a_x \equiv \frac{dA}{dx}$$

2.1.1 Sei $s(\xi,p)$ die (spezifische) Sensitivität des Rezeptors r_ξ hinsichtlich des Parameters p, und sei ρ_ξ seine Reaktionsaktivität:

$$s(\xi,p) = k\,\frac{d\rho}{da} = k\,\left(\frac{d\rho}{dp}\right)_\xi \Big/ \left(\frac{da}{dp}\right)_\xi$$

wobei

$$\int_0^\infty s(\xi,p)dp = 1\ ,$$

und k eine Normalisierungskonstante ist.

2.1.2 Angenommen, x und ξ fallen zusammen. Die Reaktion ρ_ξ des Rezeptors r_ξ auf seinen Stimulus S_ξ ist dann

$$\rho_\xi = \int_0^\infty S(\xi,p) \cdot s(\xi,p)dp = F(a_\xi)\ .$$

2.2 Dieser Ausdruck zeigt, daß weder die Modalität des Agens, noch seine parametrischen Charakteristika, noch auch der Bezug auf den Umweltpunkt x in der Reaktion des Rezeptors enkodiert werden; durch die Aktivität ρ_ξ des Rezeptors r_ξ werden lediglich bestimmte Hinweise auf das Vorhandensein eines Stimulus gegeben.

2.2.1 Da alle Rezeptoren eines Organismus in der gleichen Weise reagieren, ist klar, daß Organismen nicht imstande sind, irgendwelche Vorstellungen wie z.B. die der „Mannigfaltigkeit der Eigenschaften der Umwelt" zu bilden, wenn sie nicht auf ihren eigenen Körper Bezug nehmen, indem sie die geometrische Bedeutung der Bezeichnung ξ des Rezeptors r_ξ nutzen, die besagt: „So und so viel ($\rho = \rho_1$) gibt es an *diesem* Ort auf meinem Körper ($\xi = \xi_1$)."

2.2.2 Es ist darüber hinaus klar, daß irgendwelche Vorstellungen von „sensorischer Modalität" sich nicht aus einer „sensorischen Spezifität" ergeben können, etwa aus einer Unterscheidung der Sensitivität hinsichtlich verschiedener Para-

meter p_1 und p_2, oder unterschiedlicher Sensitivitäten s_1 und s_2 mit Bezug auf denselben Parameter p, denn alle diese Unterscheidungen werden, wie der Ausdruck 2.1.2 zeigt, „hinausintegriert". Diese Vorstellungen können sich folglich nur durch eine Unterscheidung der auf den Körper orientierten Orte der Sinneswahrnehmung ξ_1 und ξ_2 ergeben. (Kneift man den kleinen Zeh des linken Fußes, so spürt man dies nicht im Gehirn, sondern im kleinen Zeh des linken Fußes. Verschiebt man einen Augapfel, indem man ihn vorsichtig auf die Seite drückt, so wird das Umweltabbild in diesem Auge relativ zum Abbild im anderen Auge verschoben.)

2.2.3 Daraus wird klar, daß alle Schlußfolgerungen hinsichtlich der Umwelt von Ω durch Operationen über die Verteilungsfunktion ρ_ξ errechnet werden müssen. (Dies zeigt außerdem, daß diese Operationen ω_{ij} in einem gewissen Sinn an verschiedene Sensitivitäten $s_i[\xi, p_j]$ gekoppelt sind.)

2.2.4 Dies wird noch deutlicher, wenn ein physikalisches Agens in der Umwelt „Aktionen auf Distanz" erzeugt.

2.2.4.1 Sei g(x,p) die Umweltverteilung der Quellen des Agens mit der parametrischen Mannigfaltigkeit (p); sei R die Entfernung zwischen jedem beliebigen Punkt x der Umwelt und einem festgelegten Punkt x_0; und sei $\Phi(R)$ die Distanzfunktion, entsprechend welcher das Agens seine Intensität verliert. Wenn darüber hinaus der Punkt ξ_0 auf dem Körper eines Organismus mit dem Punkt x_0 zusammenfällt, dann ergibt sich die Stimulusintensität für den Rezeptor $r_{\xi 0}$ wie folgt (Bild 6):

$$S(x_0,p)_{\text{Position}} = \int_{\substack{\text{senso-}\\ \text{risches}\\ \text{Feld}}} \Phi(x - x_0)\, g(x,p)\, dx\,.$$

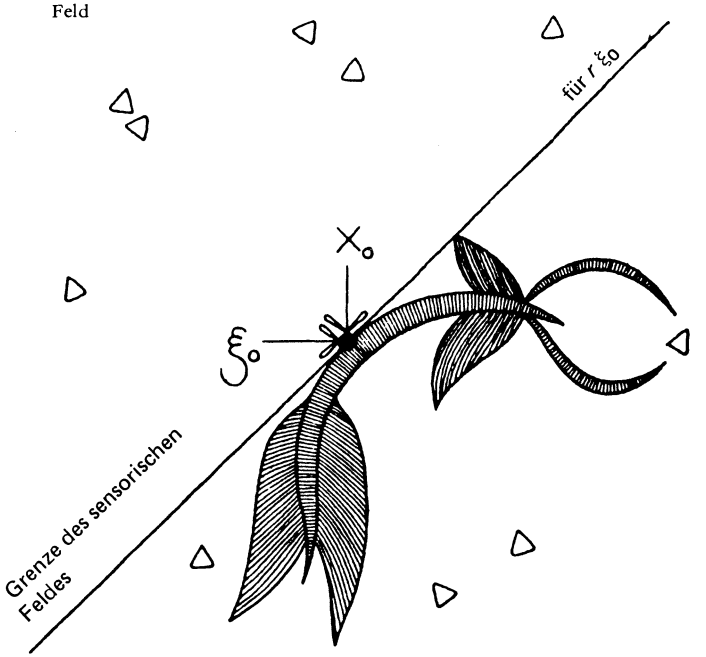


Bild 6 Geometrie des sensorischen Feldes für einen spezifischen Sensor $r_{\xi 0}$, der auf ein Agens Δ reagiert, welches über den Raum der Umwelt verteilt ist.

Und seine Reaktion ist

$$\rho_{\xi 0' \text{Position}} \equiv F(S_{\xi 0' \text{Position}})$$

(vgl. 2.1.2).

2.2.4.2 Dieser Ausdruck zeigt erneut die Unterdrückung aller raumbezogenen Hinweise mit Ausnahme der Selbstreferenz, ausgedrückt durch die körperliche Lokalisierung ξ_0 der Sinneswahrnehmung *und* durch die Position des Organismus, die durch die Grenzen des Integrals angegeben wird, welches natürlich nur über das sensorische Feld berechnet werden kann, das von der Rezeptorzelle $r_{\xi 0}$ „gesehen" wird (Bild 6).

2.2.5 Da ρ_ξ keine Hinweise auf die Art des Stimulus (p) gibt, müssen entweder ξ, der Ursprungsort der Sinnesempfindung, oder die Operation $\omega(\rho_\xi)$, oder beide die Ausbildung einer „sensorischen Modalität" bewirken.

2.2.5.1 In bestimmten Fällen ist es möglich, die räumliche Verteilung eines Agens aus der bekannten Verteilung seiner Wirkungen über eine geschlossene Oberfläche von gegebener Gestalt zu berechnen. So hat etwa die räumliche Verteilung eines elektrischen Potentials V_x eine eindeutige Lösung durch die Berechnung der Laplaceschen Gleichung

$$\Delta V = 0$$

für die gegebenen Werte V_ξ über eine geschlossene orientierbare Oberfläche (elektrischer Fisch). Weitere Beispiele ließen sich zitieren.

2.2.5.2 In einigen anderen Fällen ist es möglich, die räumliche Verteilung eines Agens aus seinen Wirkungen auf zwei kleine, aber getrennte Bereiche des Körpers zu berechnen. So läßt sich z.B. die (euklidische, dreidimensionale) Vorstellung der „Tiefe" berechnen, indem die Diskrepanz zwischen den zwei unterschiedlichen Abbildungen der „gleichen Szene" auf der Retina der beiden Augen bei binokularen Tieren aufgelöst wird (Bild 7). Sei L(x,y) ein postretinales Netzwerk, daß die Relation „x ist links von y" berechnet. Während das rechte Auge mitteilt, das Objekt „a" befinde sich links von „b" ($L_r[a,b]$), berichtet das linke Auge in entgegengesetzter Weise, daß Objekt „b" sich links von „a" befinde ($L_l[b,a]$). Ein Netzwerk B, das den unterschiedlichen Ursprung von Signalen von Zellgruppen $\{r_\xi\}_r$ und $\{r_\xi\}_l$ auf der linken bzw. rechten Seite des Körpers des Lebewesens berücksichtigt, berechnet mit $B(L_r,L_l)$ eine neue „Dimension", nämlich die Relation B(a,b): „a ist hinter b mit Bezug auf *mich*" (das Subskript s ≡ „Ich").

2.2.5.3 Die Ergebnisse dieser Berechnungen beruhen stets auf einer (geometrischen oder anderen) Relation zwischen dem Organismus und seiner Umwelt, wie sie durch die relative Vorstellung „hinten"/„hinter" ausgedrückt wird; oder auf einer Relation zwischen dem Organismus und sich selbst, wie sie durch die absoluten Vorstellungen „*mein* linkes Auge" oder „*mein* rechtes Auge" ausgedrückt wird. Dies ist der Ursprung der „Selbstreferenz".

2.2.5.4 Es ist klar, daß die Hauptlast dieser Berechnungen in den Operationen ω liegt, die über die Verteilungsfunktionen p_ξ ablaufen.

2.2.6 Damit sich irgendwelche derartigen Operationen ausbilden können, müssen Veränderungen der Sinneswahrnehmung $\delta\rho_\xi$ mit Ursachen dieser Veränderungen, die vom Organismus kontrolliert werden, verglichen werden.

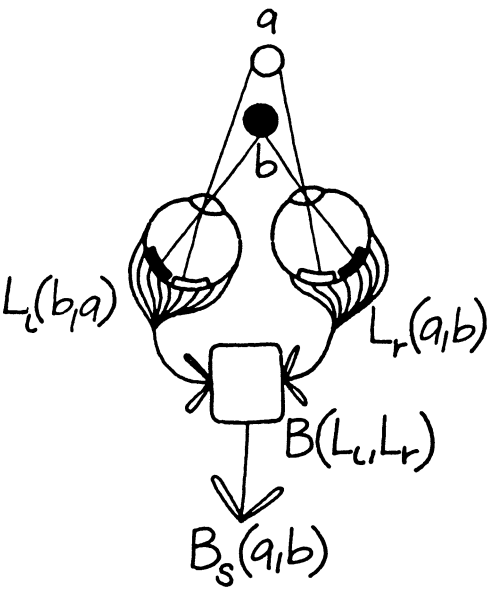


Bild 7
Berechnung der „Tiefe" durch Auflösung einer sensorischen Diskrepanz beim binokularen Sehen; (L) Netzwerke berechnen die Relation „x ist links von y"; (B) Netzwerke berechnen die Relation „x ist hinter y".

3. In einer stationären Umwelt, die hinsichtlich der Parameter p_i anisotropisch ist, verursacht eine Bewegung $\delta\Gamma$ des Organismus eine Veränderung seiner Sinneswahrnehmungen $\delta\rho_\xi$. Es ergibt sich daher

(Bewegung) → (Veränderung der Sinneswahrnehmung),

aber nicht notwendig

(Veränderung der Sinneswahrnehmung) → (Bewegung).

3.1 Die Begriffe „Veränderung der Sinneswahrnehmung" und „Bewegung" beziehen sich auf Erfahrungen des Organismus. Daher beschreibt die Notation, die ich hier verwende, diese Sachverhalte ρ_ξ, μ_ζ ausschließlich durch die Eigen-Koordinaten ξ und ζ. (μ_ζ soll die Aktivität kontraktiler Elemente m_ζ anzeigen. μ_ζ ist folglich eine äquivalente Beschreibung von $\delta\Gamma$: $\mu_\zeta \rightarrow \delta\Gamma$.)

3.1.1 Diese Begriffe sind eingeführt worden, um sie mit den Begriffen „Stimulus" und „Reaktion" eines Organismus zu konfrontieren, die sich auf die Erfahrungen desjenigen beziehen, der den Organismus *beobachtet*, nicht auf die Erfahrungen des *Organismus selbst*. Die hier verwendete Notation beschreibt diese Sachverhalte $S_x, \delta\Gamma_x$ durch die Umweltkoordinaten x. Dies ist insofern korrekt, als der Organismus Ω für einen Beobachter O ein Teil der Umwelt ist.

3.1.1.1 Damit ist klar, daß „Stimulus" nicht mit „Veränderung der Sinneswahrnehmung" und „Reaktion" nicht mit „Bewegung" gleichgesetzt werden kann, auch wenn durchaus denkbar ist, daß die komplexen Relationen, die unzweifelhaft zwischen diesen Begriffen bestehen, einmal eindeutig ermittelt werden können, wenn über den kognitiven Prozeß sowohl im Beobachter als auch im Organismus selbst mehr bekannt ist.

3.1.1.2 Aus dem unter Aussage 3 abgeleiteten *non sequitur* ergibt sich *a fortiori*:

nicht notwendig: (Stimulus) → (Reaktion).

3.2 Das Auftreten eines wahrnehmbaren Agens in schwacher Konzentration kann einen Organismus veranlassen, sich auf dieses Agens hin zu bewegen (Annäherung). Das Auftreten desselben Agens in starker Konzentration kann diesen Organismus jedoch veranlassen, sich von ihm zu entfernen (Abwendung).

3.2.1 Dies kann durch das folgende Schema abgebildet werden:

$$\begin{array}{l} S_0 \geqslant S_\xi \searrow \\ \qquad\qquad \rho_\xi \longrightarrow \omega(\rho_\xi) \\ S_0 < S_\xi \nearrow \end{array} \begin{array}{l} \nearrow \mu_\xi^+ \longrightarrow \delta\Gamma^+ \\ \\ \searrow \mu_\xi^- \longrightarrow \delta\Gamma^- \end{array}$$

(+) und (−) bezeichnen Annäherung bzw. Abwendung.

3.2.1.1 Dieses Schema repräsentiert in Minimalform

a) „Umwelt" [S]
b) „interne Repräsentation der Umwelt" ($\omega[\rho_\xi]$)
c) „Beschreibung der Umwelt" ($\delta\Gamma^+$, $\delta\Gamma^-$).

4. Die logische Struktur von Beschreibungen ergibt sich aus der logischen Struktur von Bewegungen; „Annäherung" und „Abwendung" sind die Vorläufer von „Ja" und „Nein".

4.1 Die zwei Phasen elementaren Verhaltens, „Annäherung" und „Abwendung", bilden den operationalen Ursprung der beiden fundamentalen Axiome der zweiwertigen Logik, nämlich des „Gesetzes vom ausgeschlossenen Widerspruch": $\overline{X \,\&\, \overline{X}}$ (nicht: X *und* nicht-X) und des „Gesetzes vom ausgeschlossenen Dritten": $X \vee \overline{X}$ (X *oder* nicht-X) (Bild 8).

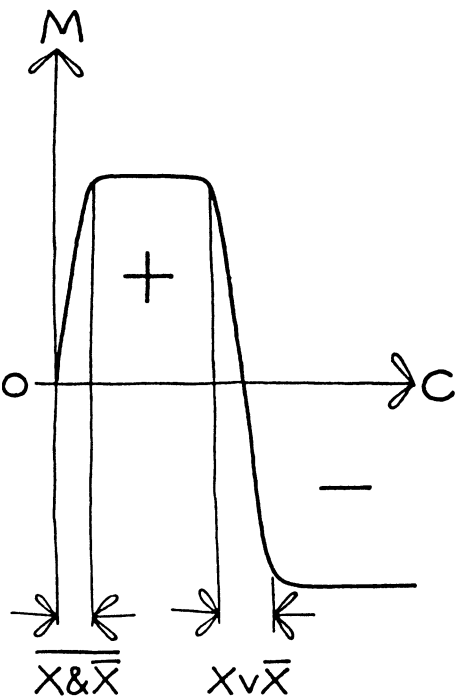


Bild 8

Die Gesetze vom „ausgeschlossenen Widerspruch" ($X \,\&\, \overline{X}$) und vom „ausgeschlossenen Dritten" ($X \vee \overline{X}$) in den Übergangszonen zwischen Bewegungslosigkeit (M = 0) und Annäherung (+), sowie zwischen Annäherung (+) und Abwendung (−), als Funktion der Konzentration (C) eines wahrnehmbaren Agens.

4.2 In Wittgensteins *Tractatus* (1961) lautet Aussage 4.0621:

„Daß aber die Zeichen ‚p' und ‚nicht-p' das gleiche sagen *können*, ist wichtig. Denn es zeigt, daß dem Zeichen ‚nicht' in der Wirklichkeit nichts entspricht. Daß in einem Satz die Verneinung vorkommt, ist noch kein Merkmal seines Sinnes (nicht-nicht-p = p)."

4.2.1 Da in der Umwelt nichts der Negation entspricht, müssen die Negation ebenso wie alle anderen „logischen Partikel" (Inklusion, Alternation, Implikation usw.) innerhalb des Organismus aufgrund seiner Wahrnehmung der Relation zwischen sich selbst und seiner Umwelt entstehen.

4.3 Beschreibungen können nicht nur im logischen Sinne bejahend oder verneinend, sondern auch wahr oder falsch sein.

4.3.1 Aus Susanne K. Langers *Philosophy in a New Key* (1951) zitieren wir:

„Der Gebrauch von Zeichen ist das allererste Zeugnis des Bewußtseins. Er ergibt sich ebenso früh in der biologischen Geschichte wie der berühmte ‚bedingte Reflex', bei dem ein begleitendes Merkmal eines Stimulus die Stimulusfunktion übernimmt. Das begleitende Merkmal wird zu einem *Zeichen* für die Situation, der die Reaktion genau angemessen ist. Dies ist der eigentliche Anfang des Denkens, denn hier liegt der Ursprung des *Irrtums* und somit der *Wahrheit*."

4.3.2 Damit ist nicht nur die logische Struktur, sondern auch der Wahrheitswert von Beschreibungen an Bewegung gekoppelt.

4.4 Die Bewegung, $\delta\Gamma$, wird intern durch Operationen über periphere Signale repräsentiert, die erzeugt werden durch

(a) Propriozeptoren p_ξ: $\quad \delta\Gamma \rightarrow \pi_\xi \rightarrow \omega(\pi_\xi)$,

(b) Sensoren r_ξ: $\quad \delta\Gamma \rightarrow \delta\rho_\xi \rightarrow \omega(\rho_\xi)$;

und Bewegung wird ausgelöst durch Operationen über die Aktivität ν_ξ zentraler Elemente n_ξ,

(c) Absicht: $\quad \omega(\nu_\xi) \rightarrow \mu_\xi \rightarrow \delta\Gamma$.

4.4.1 Da periphere Aktivität zentrale Aktivität bedingt,

$$\rho_\xi \rightarrow \nu_\xi \leftarrow \pi_\xi ,$$

ergibt sich

$$\begin{array}{ccc} \omega(\nu_\xi) & \longrightarrow & \delta\Gamma \\ \uparrow & & \downarrow \\ \delta\Gamma & \longleftarrow & \omega(\nu_\xi) \end{array} .$$

4.4.1.1 Damit zeigt sich, daß eine Konzeptualisierung von Beschreibungen der (internen Repräsentation von) Umwelt aus der Konzeptualisierung potentieller Bewegungen entsteht. Dies führt zu Ausdrücken der Form

$$\omega^{(n)}(\delta\Gamma_1, \omega^{(n-1)}[\delta\Gamma_2, \omega^{(n-2)}(\ldots[\rho_\xi])]) ,$$

d.h. zu „Beschreibungen von Beschreibungen von Beschreibungen...", oder äquivalent, zu „Repräsentationen von Repräsentationen von Repräsentationen... ."

5. Die Information eines Ereignisse E besteht in der Ausbildung von Operationen ω, die die interne Repräsentation $\omega(\rho_\xi)$ oder die Beschreibung $\delta\Gamma$ dieses Ereignisses kontrollieren.

5.1 Das Maß der möglichen Auswahl von Repräsentationen ($\omega_i[E]$) oder Beschreibungen ($\delta\Gamma_i[E]$) dieses Ereignisses – oder der Wahrscheinlichkeiten p_i seines

Auftretens – ist der „Informationsbetrag" dieses Ereignisses mit Bezug auf den Organismus Ω. ($H[E,\Omega] = -\log_2 p_i$, d.h. der negative Mittelwert[1]) aller $[\log_2 p_i]$.)

5.1.1 Damit ist gezeigt, daß Information ein relativer Begriff ist. Dies gilt auch für H.

5.2 Die Klasse verschiedener Repräsentationen eines Ereignisses E legt eine Äquivalenzklasse für verschiedene Ereignisse $(E_i[\omega]) \equiv E$ fest. Damit ist ein Maß der Anzahl von Ereignissen (E_i), die eine kognitive Einheit, eine „Kategorie E" (Quastler 1958) bilden – oder der Wahrscheinlichkeiten p_i ihres Auftretens – wiederum der „Informationsbetrag", H, der von einem Beobachter aufgrund seiner Wahrnehmung des Auftretens eines dieser Ereignisse empfangen wird.

5.2.1 Dies bedeutet, daß der Informationsbetrag eine Zahl ist, die von der Wahl einer Kategorie, d.h. einer kognitiven Einheit, abhängt.

5.3 Ich zitiere aus einer Arbeit von Jerzy Konorski (1962):

> „Es ist nicht so, wie wir aufgrund unserer Introspektion geneigt sind anzunehmen, daß der Empfang von Information und deren Nutzung zwei getrennte Prozesse sind, die auf beliebige Art miteinander kombiniert werden können; im Gegenteil, Information und ihre Nutzung sind untrennbar und bilden in Wirklichkeit einen einzigen Prozeß."

5.3.1 Diese Prozesse sind die Operationen ω, und sie werden durch die strukturelle und funktionale Organisation nervöser Aktivität verwirklicht.

5.4 Seien ν_i die Signale, die die einzelnen Fasern i durchlaufen, und sei $\nu^{(1)}$ das Ergebnis einer Interaktion von N Fasern (i = 1,2,...,N):

$$\nu^{(1)} = F^{(1)}(\nu_1, \nu_2, \ldots, \nu_N) \equiv F^{(1)}([\nu_i]) .$$

5.4.1 Es ist nützlich, die Aktivität einer Teilmenge dieser Fasern als entscheidend für die funktionale Interaktion der übrigen Fasern aufzufassen („Inhibition" verändert die funktionale Interaktion „exzitatorischer" Signale). Dies läßt sich durch einen Formalismus ausdrücken, der die durch die übrigen Fasern errechneten Funktionen bestimmt:

$$\nu^{(1)} = f^{(1)}_{[\nu_j]}([\nu_i]) , \quad j \neq i .$$

Die Übereinstimmung zwischen den Werten v des Zeilenvektors ν_j und den entsprechenden Funktionen $f^{(1)}_v$ bildet ein Funktional für die Klasse der Funktionen $f^{(1)}_{[\nu_j]}$.

5.4.1.1 Diese Notation macht klar, daß die Signale selbst zum Teil verantwortlich sind für die Festlegung der über ihnen ausgeführten Operationen.

5.4.2 Die Abbildung, die diese Übereinstimmung herstellt, wird gewöhnlich als „strukturelle Organisation" dieser Operationen interpretiert, während die so erzeugte Menge von Funktionen ihre „funktionale Organisation" ist.

5.4.2.1 Damit wird deutlich, daß die Unterscheidung zwischen der strukturellen und funktionalen Organisation kognitiver Prozesse vom Standpunkt des Beobachters abhängt.

5.4.3 Betrachtet man N Fasern, dann gibt es zwei 2^N mögliche Interpretationen (die Menge aller Teilmengen von N) der funktionalen und strukturellen Organisa-

---

[1]) Der Mittelwert einer Menge von Größen x_j, deren Auftretenswahrscheinlichkeit von p_j durch $\bar{x}_j = \Sigma x_j \cdot p_j$ gegeben ist.

tion derartiger Operationen. Sind alle Interpretationen gleichwahrscheinlich, dann ist die „Unsicherheit" dieses Systems hinsichtlich seiner Interpretierbarkeit $H = \log_2 2^N = N$ bits.

5.5 Seien $\nu_i^{(1)}$ die Signale, die die einzelnen Fasern i durchlaufen, und sei $\nu^{(2)}$ das Ergebnis einer Interaktion von N_1 solcher Fasern ($i = 1,2,\ldots,N_1$):

$$\nu^{(2)} = F^{(2)}([\nu_i^{(1)}])$$

oder rekursiv aus 5.4:

$$\nu^{(k)} = F^{(k)}(F^{[k-1]}[F^{(k-2)}(\ldots F^{[1]}[\nu_i])]) .$$

5.5.1 Da die $F^{(k)}$ als Funktionale $f^{(k)}_{[\nu_i]}$ interpretiert werden können, erreichen wir damit einen Kalkül rekursiver Funktionale für die Repräsentation kognitiver Prozesse ω.

5.5.1.1 Dies wird besonders deutlich, wenn $\nu_i^{(k-t)}$ die Aktivität der Faser i in dem Zeitintervall t vor ihrer gegenwärtigen Aktivität $\nu_i^{(k)}$ bezeichnet, d.h. die Rekursion in 5.5 als eine Rekursion in der Zeit interpretiert werden kann.

5.5.2 Der in 5.5 zur Repräsentation kognitiver Prozesse, ω, eingesetzte Formalismus rekursiver Funktionale ist der lexikalischen Definitionsstruktur von Substantiven isomorph. Im wesentlichen bezeichnet ein Substantiv eine Klasse $cl^{(1)}$ von Gegenständen. Durch eine Definition wird es als Element einer umfassenderen Klasse $cl^{(2)}$ ausgewiesen, die ihrerseits durch ein Substantiv benannt wird, welches seinerseits durch eine Definition als Mitglied einer noch umfassenderen Klasse $cl^{(3)}$ ausgewiesen wird, usw. [Fasan → Vogel → Tier → Organismus → Gegenstand]:

$$cl^{(n)} = (cl^{[n-1]}_{i_{n-1}}\,[cl^{[n-2]}_{i_{n-2}}(\ldots[cl^{(1)}_{i_1}])]) .$$

Dabei vertritt die Notation (e_i) eine aus Elementen e_i gebildete Klasse, und subskribierte Subskripte werden benützt, um diese Subskripte mit den entsprechenden Superskripten zu verknüpfen.

5.5.2.1 Das n* höchster Ordnung in dieser Hierarchie von Klassen wird immer durch einen einzigen undefinierten Begriff repräsentiert, etwa durch „Gegenstand", „Entität", „Akt" usw., der sich auf Grundvorstellungen der Wahrnehmungsfähigkeit schlechthin bezieht.

5.6 Kognitive Prozesse schaffen Beschreibungen der Umwelt, d.h. Information über die Umwelt.

6. Die Umwelt enthält keine Information. Die Umwelt ist wie sie ist.

Danksagung

Humberto Maturana, Gotthard Günther und Ross Ashby danke ich sehr herzlich für ihr unermüdliches Bemühen, mich in Fragen des Lebens, der Logik und großer Systeme aufzuklären, Lebbeus Woods für Zeichnungen, die meine Behauptungen besser verdeutlichen als ich es mit Worten allein könnte. Für alle verbleibenden Mängel meiner Analyse und Darstellung trage aber ich allein die Verantwortung und keiner dieser meiner Freunde, die mir so großzügig ihre Zeit geschenkt haben.

Über selbst-organisierende Systeme und ihre Umwelten

Über selbst-organisierende Systeme und ihre Umwelten *

Ich zögere ein wenig mit den einleitenden Bemerkungen meines Vortrages, da ich fürchte, die Gefühle gerade jener zu verletzen, die diese Tagung über selbst-organisierende Systeme so großzügig unterstützt haben. Andererseits glaube ich aber, eine Antwort auf die Frage vorschlagen zu können, die Dr. Weyl in seiner ebenso sachgerechten wie anregenden Einführung gestellt hat: „Was macht ein selbst-organisierendes System aus?" Ich hoffe also, daß Sie mir vergeben werden, wenn ich meine Ausführungen mit der folgenden These eröffne: „Es gibt keinerlei Systeme, die sich selbst organisieren!"

Angesichts des Themas dieser Tagung habe ich nun einen sehr überzeugenden Beweis für diese These zu liefern! Dies wird allerdings nicht allzu schwierig sein, sollte der geheime Zweck dieser Versammlung nicht gerade in einer Verschwörung bestehen, den Zweiten Hauptsatz der Thermodynamik abzuschaffen. Ich werde also nun die Nicht-Existenz selbst-organisierender Systeme durch eine *reductio ad absurdum* der Annahme beweisen, daß es so etwas gibt wie ein System, das sich selbst organisiert.

Nehmen wir ein finites Universum, U_0, an, so klein oder so groß, wie sie es wünschen (vgl. Bild 1a), das in eine adiabatische Schale eingeschlossen ist, welche dieses finite Universum von dem „Meta-Universum" trennt, in das es eingebettet ist. Nehmen wir des weiteren an, daß es in diesem Universum U_0 eine geschlossene Oberfläche gibt, die es in zwei voneinander völlig getrennte Teile teilt: der eine Teil wird vollständig durch ein selbst-organisierendes System S_0 besetzt, während

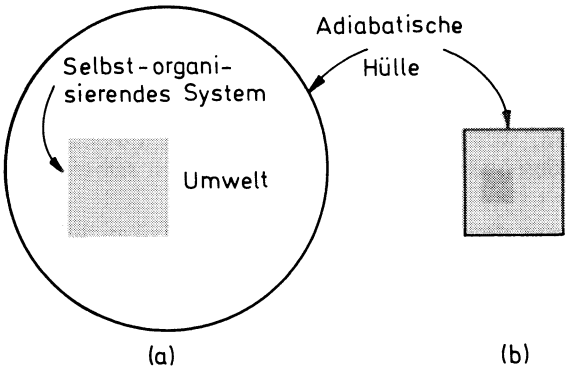


Bild 1

* Überarbeitete Fassung eines Vortrags auf dem "Interdisciplinary Symposium on Self-Organizing Systems", Chicago/Illinois, vom 5. Mai 1960.

wir den anderen Teil die Umwelt E_0 dieses selbst-organisierenden Systems nennen können:

$$S_0 \,\&\, E_0 = U_0.$$

Ich darf anfügen, daß es irrelevant ist, ob sich unser selbst-organisierendes System innerhalb oder außerhalb der geschlossenen Oberfläche befindet. In Bild 1 liegt das System innerhalb.

Wird diesem selbst-organisierenden System nun Gelegenheit gegeben, sich eine Zeitlang damit zu beschäftigen, sich selbst zu organisieren, dann wird seine Entropie während dieser Zeit unzweifelhaft abnehmen:

$$\frac{\delta S_s}{\delta t} < 0,$$

sonst würden wir es ja auch nicht als selbst-organisierendes System bezeichnen, sondern lediglich als mechanisches $\delta S_s/\delta t = 0$ oder als thermodynamisches $\delta S_s/\delta t > 0$ System. Die Entropie im übrigen Teil unseres finiten Universums, d.h. die Entropie der Umwelt, muß also zunehmen,

$$\frac{\delta S_E}{\delta t} > 0,$$

sonst würde der Zweite Hauptsatz der Thermodynamik verletzt. Wenn nun einige der Prozesse, die zur Abnahme der Entropie des Systems beigetragen haben, irreversibel sind, dann stellen wir fest, daß die Entropie des Universums U_0 höher ist als zu dem Zeitpunkt, da unser System begann, sich selbst zu organisieren. Der Zustand des Universums weist folglich weniger Organisation auf als vor $\delta S_U/\delta t > 0$, mit anderen Worten, die Aktivität des Systems ist des-organisierend, und wir können ein derartiges System zu Recht als „des-organisierendes System" bezeichnen.

Man muß nun aber feststellen, daß es unfair ist, das System für Veränderungen im gesamten Universum verantwortlich zu machen, und daß diese offenbare Inkonsistenz dadurch entstand, daß wir nicht nur das System selbst betrachtet, sondern auch die Umwelt des Systems in unsere Überlegung einbezogen haben. Wenn man eine zu große adiabatische Hülle konstruiert, führt man Prozesse ein, die für unsere Überlegungen überhaupt nicht von Bedeutung sind. Also gut, lassen wir die adiabatische Hülle mit der geschlossenen Oberfläche zusammenfallen, die das System vorhin von seiner Umwelt getrennt hat (Bild 1b). Dieser Schritt räumt nicht nur den obigen Einwand aus, er erlaubt mir außerdem zu zeigen, daß ein selbst-organisierendes System, welches von dieser Hülle umschlossen wird, sich nicht nur als ein des-organisierendes, sondern sogar als ein sich selbst des-organisierendes System erweist.

Mein früheres Beispiel mit der großen Hülle macht klar, daß auch hier – sollten irreversible Prozesse auftreten – die Entropie des Systems, das nunmehr innerhalb der Hülle liegt, zunehmen muß, daß daher im Laufe der Zeit das System sich selbst des-organisieren würde, obwohl die Entropie in gewissen Bereichen tatsächlich abgenommen haben mag. Man kann nun darauf hinweisen, daß wir unsere Hülle eben nur um diese Bereiche hätten wickeln sollen, da sie der eigentliche selbst-organisierende Teil unseres Systems zu sein scheinen. Aber ich könnte hier erneut dasselbe Argument wie früher anbringen, diesmal nur mit Bezug auf einen kleineren Bereich, und so könnten wir für immer weiter fortfahren, bis unser an-

geblich selbst-organisierendes System in die glücklichen ewigen Jagdgründe des Infinitesimalen verschwunden ist.

Obwohl ich nun diesen Beweis der Nicht-Existenz selbst-organisierender Systeme vorgeschlagen habe, möchte ich den Begriff „selbst-organisierendes System" weiterhin verwenden. Dabei muß allerdings besonders beachtet werden, daß dieser Begriff sinnlos wird, wenn das System in engem Kontakt mit einer Umwelt steht, *die verfügbare Energie und Ordnung besitzt*, und mit der unser System durch ständige Interaktion verbunden ist, so daß es in bestimmter Weise auf Kosten dieser Umwelt zu „leben" vermag.

Obwohl ich nun nicht im Detail die interessante Diskussion des Energieflusses von der Umwelt in das System und zurück aufnehmen werde, möchte ich kurz die beiden Denktraditionen erwähnen, die mit diesem Problem zusammenhängen. Die eine sieht Energiefluß und Signalfluß als eine eng verknüpfte Ein-Kanal-Angelegenheit an (d.h. die Botschaft trägt auch die Nahrung, oder wenn Sie wollen, Signal und Nahrung sind synonym), während die andere diese beiden Größen sorgfältig voneinander trennt, auch wenn es in dieser Theorie einen signifikanten Zusammenhang zwischen Signalfluß und Energievorrat gibt.

Ich bekenne, daß ich der zweiten Denkweise anhänge, und ich bin besonders glücklich, daß im weiteren Verlauf dieser Tagung Gordon Pask diese Sichtweise in seinem Vortrag „Die Naturgeschichte von Netzwerken" (1960) viel klarer darstellen wird, als ich dies jemals vermöchte.

Was mich im gegenwärtigen Augenblick besonders interessiert, das ist nicht so sehr die Energie aus der Umwelt, die vom System verdaut wird, sondern die Nutzung der Umweltordnung durch das System. Mit anderen Worten, die Frage, die ich gerne beantworten würde, lautet: „Wieviel an Ordnung, wenn überhaupt, kann unser System von der Umwelt assimilieren?"

Bevor ich diese Frage anpacke, muß ich noch zwei Hürden nehmen, die jeweils Probleme im Zusammenhang mit der Umwelt darstellen. Da Sie ohne Zweifel bemerkt haben, daß in meiner Philosophie selbst-organisierender Systeme die Umwelt solcher Systeme eine *conditio sine qua non* darstellt, bin ich zunächst verpflichtet zu zeigen, in welchem Sinne wir über die Existenz einer solchen Umwelt reden können. Sodann habe ich zu zeigen, daß eine solche Umwelt strukturiert sein muß, wenn sie existiert.

Das erste Problem, das ich nun zu eliminieren beabsichtige, ist vielleicht eines der ältesten philosophischen Probleme, mit dem die Menschheit hat leben müssen. Dieses Problem entsteht dann, wenn wir, die Menschen, uns selbst als selbst-organisierende Systeme auffassen. Wir können dann darauf bestehen, daß uns Introspektion nicht gestattet zu entscheiden, ob die Welt, wie wir sie sehen, „real" oder bloß eine Phantasmagorie, ein Traum, eine Illusion unserer Einbildungskraft ist. Eine Entscheidung dieses Dilemmas ist insofern wichtig für meine Diskussion, als für den Fall, daß die letztere Alternative gilt, meine ursprüngliche These, die die Unsinnigkeit der Vorstellung eines isolierten selbst-organisierenden Systems behauptet, kläglich in sich zusammenbrechen müßte.

Ich möchte nun versuchen, die Realität der Welt, die wir wahrnehmen, zu zeigen, und zwar durch eine *reductio ad absurdum* der These: Diese Welt existiert nur in unserer Vorstellung und die einzige Realität ist das sich etwas vorstellende „Ich".

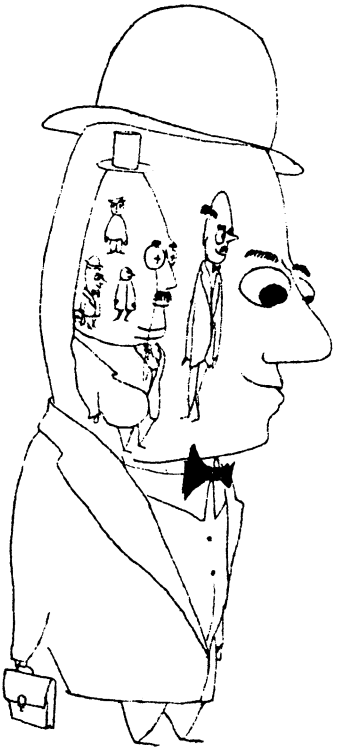


Bild 2

Dank des künstlerischen Beistandes von Gordon Pask, der diese und einige meiner weiteren Behauptungen so wunderschön illustriert hat (Bilder 2, 5, 6), wird es für mich leicht, meine Argumentation zu entwickeln.

Nehmen wir für den Augenblick an, daß ich der erfolgreiche Geschäftsmann mit der Melone auf dem Kopf in Bild 2 bin und daß ich darauf bestehe, die einzige Realität zu sein, während alles übrige nur in meiner Vorstellung auftritt. Ich kann nun aber nicht leugnen, daß in meiner Vorstellung Menschen, Naturwissenschaftler, andere erfolgreiche Geschäftsleute usw. erscheinen, wie z.B. bei dieser Tagung hier. Da ich diese Erscheinungen in vielen Hinsichten als mir selber ähnlich erkenne, muß ich ihnen auch das Privileg zugestehen, daß sie selber darauf bestehen, daß sie die einzige Realität sind und alles andere nur ein Gebräu ihrer Einbildung darstellt. Andererseits können sie wiederum nicht leugnen, daß ihre eigene Phantasie von Menschen bevölkert ist ... – und einer dieser Menschen könnte ich sein, mit der Melone auf dem Kopf und allem Drum und Dran!

Damit haben wir den Zirkel unseres Widerspruchs geschlossen: Wenn ich annehme, daß ich die einzige Realität bin, dann stellt sich heraus, daß ich nur die Vorstellung von jemand anders bin, der seinerseits annimmt, daß *er* die einzige Realität ist. Natürlich läßt sich dieses Paradox leicht dadurch auflösen, daß man die Realität der Welt postuliert, in der wir alle so glücklich blühen und gedeihen.

Nachdem wir nun die Realität wieder eingesetzt haben, ist die interessante Feststellung zu machen, daß die Realität als konsistenter Bezugsrahmen für zumindest zwei Beobachter auftritt. Dies wird besonders klar dann, wenn erkannt wird, daß mein „Beweis" exakt nach dem „Relativitätsprinzip" modelliert war, das im wesentlichen folgendes besagt: Wenn eine Hypothese, die auf eine Menge von Gegenständen anwendbar ist, für einen Gegenstand und noch für einen anderen Gegenstand gilt, dann gilt sie für beide Gegenstände gleich und muß schließlich für alle Gegenstände der Menge akzeptiert werden. In der Ausdrucksweise der symbolischen Logik gilt daher:

$$(\exists x)\,[H(a) \,\&\, H(x) \rightarrow H(a + x)] \rightarrow (x)\,H(x). \qquad (1)$$

Kopernikus etwa hätte diese Argumentation auf folgende Weise zu seinem Vorteil verwenden können: Wenn wir an ein geozentrisches System glauben, [H(a)], dann könnten z.B. die Venusianer an ein venuzentrisches System glauben, [(Hx)]. Da wir aber nicht beides sein können, Zentrum und Epizykloid zur gleichen Zeit, [H(a + x)], kann mit einem planetozentrischen System etwas nicht stimmen.

Man sollte jedoch nicht übersehen, daß der oben formulierte Ausdruck $\mathscr{R}$ (H) keine Tautologie ist, also eine empirisch gehaltvolle Aussage sein muß.[1] Diese Aussage schafft eine Möglichkeit, über die Existenz einer Umwelt zu reden.

Bevor ich nun zu meiner ursprünglichen Frage zurückkehren kann, wieviel Ordnung ein selbst-organisierendes System aus seiner Umwelt assimilieren kann, muß ich zeigen, daß unsere Umwelt eine gewisse Struktur aufweist. Dies läßt sich in der Tat sehr leicht allein durch die Feststellung tun, daß wir uns ganz augenscheinlich noch nicht in dem schrecklichen Zustand des Boltzmannschen „Wärmetods" befinden. Daher nimmt die Entropie immer noch zu, und dies bedeutet, daß es eine gewisse Ordnung geben *muß* – zumindest jetzt –, andernfalls könnten wir Ordnung nicht verlieren.

Lassen Sie mich die bisher vorgelegten Argumente kurz zusammenfassen:

1. Mit einem selbst-organisierenden System meine ich jenen Teil eines Systems, der sich Energie und Ordnung aus seiner Umwelt einverleibt.
2. Es gibt die Realität der Umwelt in dem Sinne, den die Akzeptanz des Relativitätsprinzips nahelegt.
3. Die Umwelt besitzt Struktur.

Wir wollen uns nun wieder unseren selbst-organisierenden Systemen zuwenden. Von solchen Systemen erwarten wir, daß sie ihre innere Ordnung vergrößern. Um diesen Prozeß zu beschreiben, wäre es zuerst einmal schön, wenn wir „innere" definieren könnten, und zweitens, wenn wir irgendein Maß für Ordnung hätten. Das erste Problem entsteht immer dann, wenn wir es mit Systemen zu tun haben, die nicht von einer Haut überwachsen sind. In solchen Fällen liegt es an uns, die geschlossene Grenze unseres Systems zu definieren. Dies kann jedoch Schwierigkeiten bereiten, denn wenn wir einen bestimmten Bereich im Raum als den intuitiv richtigen Platz für unser selbst-organisierendes System ansehen, dann kann

---

[1] Dies hat Wittgenstein bemerkt, obwohl seine Einsicht das Prinzip der mathematischen Induktion betraf. Der enge Zusammenhang aber zwischen dem Induktions- und dem Relativitätsprinzip ist durchaus evident. Ich würde sogar die Behauptung wagen, daß das Prinzip der mathematischen Induktion das Relativitätsprinzip der Zahlentheorie ist.

es sich herausstellen, daß eben dieser Bereich überhaupt keine selbst-organisierenden Eigenschaften aufweist, und daß wir gezwungen sind, eine andere Wahl zu treffen, in der Hoffnung, diesmal größeres Glück zu haben. Genau diese Schwierigkeit tritt etwa auf in Verbindung mit dem Problem der „Lokalisierung von Funktionen" im Großhirn.

Wir können die Argumentation andersherum aufziehen, indem wir unsere Grenze jeweils als die Hülle desjenigen Bereichs im Raum definieren, der die gewünschte Zunahme an Ordnung zeigt. Aber auch damit geraten wir in Schwierigkeiten; ich kenne nämlich kein Gerät, das angeben würde, ob es an einen selbst-*des*organisierenden oder einen selbst-organisierenden Bereich angeschlossen ist, und das uns somit eine solide operationale Definition liefert.

Eine weitere Schwierigkeit ergibt sich aus der Möglichkeit, daß diese selbst-organisierenden Bereiche sich nicht nur ständig im Raum bewegen und in ihrer Form verändern, sondern daß sie auch hier und dort spontan auftreten und wieder verschwinden können, so daß der „Ordometer" diesen völlig ungreifbaren Systemen nicht nur nachlaufen, sondern auch den Ort ihrer Entstehung riechen muß!

Mit dieser kleinen Abschweifung wollte ich lediglich darauf hinweisen, daß wir sehr vorsichtig sein müssen, wenn wir in diesem Zusammenhang das Wort „innen" verwenden, da ein Beobachter trotz der Tatsache, daß seine Position festgelegt worden ist, beträchtliche Schwierigkeiten haben kann, das auszudrücken, was er sieht.

Wir wollen uns nun dem anderen Punkt zuwenden, den ich erwähnt habe, nämlich dem Versuch, ein angemessenes Maß für Ordnung zu finden. Ich persönlich bin der Auffassung, daß wir mit diesem Begriff zwei Sachverhalte beschreiben möchten. Erstens könnten wir daran interessiert sein, bestimmte Beziehungen zwischen den Elementen einer Menge zu erklären, die die möglichen Anordnungen der Elemente dieses Systems Einschränkungen unterwerfen. In dem Maße, in dem die Organisation des Systems zunimmt, werden immer mehr dieser Beziehungen erkennbar. Zweitens scheint mir Ordnung eher eine relative denn eine absolute Konnotation zu haben; der Begriff ist nämlich bezogen auf die maximale Unordnung der Elemente der Menge. Daraus ergibt sich, daß es bequem wäre, wenn das Ordnungsmaß Werte zwischen 0 und 1 annehmen würde, um im ersten Fall maximale Unordnung, und im zweiten maximale Ordnung auszudrücken. Damit wird die „Negentropie" als Ordnungsmaß eliminiert, da Negentropie für Systeme in vollständiger Unordnung stets finite Werte annimmt. Was Shannon (1949) aber als „Redundanz" definiert hat, scheint mir maßgeschneidert für die Beschreibung der Ordnung, wie ich sie mir vorstelle. Nach Shannons Definition der Redundanz ergibt sich:

$$R = 1 - \frac{H}{H_m}, \qquad (2)$$

wobei H/H_m das Verhältnis zwischen der Entropie H einer Informationsquelle und dem Maximumwert H_m ist, den sie erreichen könnte, solange sie auf dieselben Symbole beschränkt bleibt. Shannon nennt dieses Verhältnis die „relative Entropie". Dieser Ausdruck erfüllt ganz offensichtlich die Anforderungen an ein Maß für Ordnung, wie ich sie vorhin aufgezählt habe. Wenn sich das System im Zustand maximaler Unordnung $H = H_m$ befindet, dann ist R gleich 0; wenn die Elemente des Systems dagegen so geordnet sind, daß mit Gegebensein eines Elements die

Position aller anderen Elemente determiniert ist, dann verschwindet die Entropie – bzw. der Grad der Unsicherheit –, R wird 1 und zeigt damit vollkommene Ordnung an.

Natürlich erwarten wir von einem selbst-organisierenden System, daß die Ordnung des Systems, wie sie in einem bestimmten Anfangszustand gegeben ist, im Laufe der Zeit zunimmt. Mit unserem Ausdruck (2) können wir sogleich das Kriterium dafür angeben, daß ein System sich selbst organisiert, d.h. daß die Rate der Veränderung von R positiv ist:

$$\frac{\delta R}{\delta t} > 0. \tag{3}$$

Wenn wir die Gleichung (2) mit Bezug auf die Zeit differenzieren und die Ungleichung (3) hinzuziehen, ergibt sich:

$$\frac{\delta R}{\delta t} = -\frac{H_m(\delta H/\delta t) - H(\delta H_m/\delta t)}{H_m{}^2} \tag{4}$$

Da $H_m{}^2 > 0$ unter allen Bedingungen gilt (es sei denn, wir beginnen mit Systemen, die sich nur als ständig vollkommen geordnet denken lassen: $H_m = 0$), können wir die Bedingung dafür, daß ein System sich selbst organisiert, mithilfe des Entropiekonzepts ausdrücken:

$$H\frac{\delta H_m}{\delta t} > H_m\frac{\delta H}{\delta t}. \tag{5}$$

Um die Bedeutung dieser Gleichung klar zu machen, möchte ich zuerst zwei besondere Fälle kurz erörtern, jene Fälle nämlich, in denen jeweils einer der beiden Ausdrücke H bzw. H_m als konstant angenommen wird.

(a) H_m = konstant

Wir wollen zunächst den Fall betrachten, in dem H_m, d.h. die maximale Entropie des Systems, konstant bleibt, denn dies ist der Fall, den man sich gewöhnlich vorstellt, wenn man von sich selbst organisierenden Systemen spricht. Wenn H_m als konstant angenommen wird, dann verschwindet die zeitabhängige Ableitung von H_m, und es ergibt sich aus Gleichung (5):

$$\frac{\delta H_m}{\delta t} = 0 \ldots \frac{\delta H}{\delta t} < 0. \tag{6}$$

Diese Gleichung stellt ganz einfach fest, daß die Entropie des Systems im Laufe der Zeit abnimmt. Das wußten wir bereits, – wir können uns nun aber fragen, wie dies erreicht werden kann? Da die Entropie des Systems von der Wahrscheinlichkeitsverteilung der in bestimmten unterscheidbaren Zuständen befindlichen Elemente abhängt, muß sich diese Wahrscheinlichkeitsverteilung klarerweise verändern, wenn H reduziert werden soll. Wir können uns dies ebenso wie seine Verwirklichung verdeutlichen, indem wir auf die Faktoren achten, die die Wahrscheinlichkeitsverteilung bestimmen. Einer dieser Faktoren könnte in bestimmten Eigenschaften unserer Elemente liegen, die es mehr oder minder wahrscheinlich machen, daß ein Element in einem bestimmten Zustand anzutreffen sein wird. Nehmen wir z.B. an, daß der entsprechende Zustand darin besteht, „in

einem Loch einer bestimmten Größe zu sein". Die Wahrscheinlichkeit dafür, daß Elemente sich in diesem Zustand finden, die größer sind als das Loch, ist klarerweise Null. Wenn die Elemente also wie kleine Ballons langsam aufgeblasen werden, wird sich die Wahrscheinlichkeitsverteilung ständig ändern. Ein weiterer Faktor, der die Wahrscheinlichkeitsverteilung beeinflußt, könnte in gewissen anderen Eigenschaften unserer Elemente liegen, die die bedingten Wahrscheinlichkeiten dafür bestimmen, daß ein Element in bestimmten Zuständen angetroffen wird, wenn der Zustand anderer Elemente des Systems gegeben ist. Wiederum wird eine Veränderung dieser bedingten Wahrscheinlichkeiten die Wahrscheinlichkeitsverteilung und somit die Entropie des Systems verändern. Da alle diese Veränderungen im Inneren stattfinden, will ich einen „inneren Dämon" für diese Veränderungen verantwortlich machen. Er bewirkt, daß H abnimmt, z.B. dadurch, daß er eifrig die kleinen Ballons aufbläst und somit die Wahrscheinlichkeitsverteilung verändert, oder daß er die bedingten Wahrscheinlichkeiten verschiebt, indem er Verbindungen zwischen Elementen herstellt. Da uns die Aufgaben dieses Dämons ziemlich vertraut sind, möchte ich mich einen Augenblick von ihm ab- und einem anderen zuwenden und den zweiten Spezialfall erörtern, den ich bereits erwähnt habe, jenen nämlich, in dem H als konstant angenommen wird.

(b) H = konstant

Wird die Entropie des Systems als konstant angenommen, dann verschwindet ihre zeitabhängige Ableitung und es ergibt sich aus Gleichung (5):

$$\frac{\delta H}{\delta t} = 0 \ldots \frac{\delta H_m}{\delta t} > 0. \tag{7}$$

Es ergibt sich damit das eigentümliche Resultat, daß wir nach unserer früher gegebenen Definition von Ordnung ein selbst-organisierendes System vor uns haben können, wenn die maximale Unordnung des Systems zunimmt. Nun scheint dies auf den ersten Blick eine ziemlich triviale Angelegenheit zu sein, da man sich leicht einfache Prozesse vorstellen kann, in denen diese Bedingung erfüllt ist. Betrachten wir als einfaches Beispiel ein System, das aus N Elementen zusammengesetzt ist, die bestimmte beobachtbare Zustände einnehmen können. In den meisten Fällen läßt sich eine Wahrscheinlichkeitsverteilung für die Gesamtzahl der Elemente in diesen Zuständen so berechnen, daß H maximiert und ein Ausdruck für H_m abgeleitet wird. Aufgrund der Tatsache, daß die Entropie (oder der Informationsbetrag) mit dem Logarithmus der Wahrscheinlichkeiten verknüpft ist, läßt sich ohne Schwierigkeiten zeigen, daß die Ausdrücke für H_m im allgemeinen die folgende Form haben[2]):

$$H_m = C_1 + C_2 \log_2 N.$$

Daraus ergibt sich unmittelbar ein Verfahren, H_m zu vergrößern, nämlich so, daß die Anzahl der Elemente, die das System bilden, vermehrt wird; mit anderen Worten, ein System, das durch die Einfügung neuer Elemente wächst, vergrößert seine maximale Entropie, und wir müssen mit aller nötigen Fairneß dieses System

[2]) Vgl. auch den Anhang.

als ein Mitglied der geschätzten Familie selbst-organisierender Systeme anerkennen, da es dem Kriterium hierfür (Gleichung 7) entspricht.

Wenn es nun aber schon genügt, einem System neue Elemente hinzuzufügen, um es zu einem selbst-organisierenden System zu machen, dann ließe sich behaupten, daß auch ein Eimer zu einem selbst-organisierenden System wird, wenn man Sand in ihn schüttet. Irgendwie scheint dies aber – ganz gelinde gesagt – nicht unserer intuitiven Hochachtung der Mitglieder unserer geschätzten Familie gerecht zu werden. Und dies ist auch richtig so, denn eine derartige Behauptung mißachtet die Voraussetzung, unter der diese Aussage abgeleitet wurde, nämlich die, daß im Prozeß der Hinzufügung neuer Elemente die Entropie H des Systems konstant gehalten werden muß. Im Falle des mit Sand gefüllten Eimers dürfte dies eine äußerst kitzlige Aufgabe darstellen, deren Erfüllung man sich vielleicht so vorstellen könnte, daß die neuen Teilchen hinsichtlich bestimmter unterscheidbarer Zustände, etwa Lage, Richtung usw. in genau die Ordnung gebracht werden, in der sich die Teilchen befinden, die im Augenblick der Hinzufügung der Neuankömmlinge vorhanden sind. Ganz augenscheinlich erfordert diese Aufgabe, H_m zu vergrößern und gleichzeitig H konstant zu halten, übermenschliche Geschicklichkeit, und wir können daher dafür einen weiteren Dämon anstellen, den ich den „äußeren Dämon" nennen will, und dessen Aufgabe darin besteht, in das System nur jene Elemente hineinzulassen, deren Zustand mit den Bedingungen zumindest der konstanten inneren Entropie übereinstimmt. Wie Sie nun sicherlich bereits bemerkt haben, ist dieser Dämon ein enger Verwandter des Maxwellschen Dämons, nur sind diese Burschen heutzutage leider nicht mehr so gut wie in alten Zeiten. Vor dem Jahre 1927 (Heisenberg) waren sie nämlich noch imstande, irgendein beliebiges kleines Loch zu überwachen, durch welches der Neuankömmling hindurchmußte, und mit beliebig großer Genauigkeit sein Bewegungsmoment zu prüfen. Heute sind die Dämonen, die bestimmte Löcher bewachen, leider unfähig, eine verläßliche Prüfung des Bewegungsmoments vorzunehmen und umgekehrt. Ihre Möglichkeiten sind leider Gottes durch Heisenbergs Unbestimmtheitsprinzip eingeschränkt worden.

Nachdem ich die beiden Spezialfälle erörtert habe, bei denen jeweils nur ein Dämon am Werke ist, während der andere angekettet bleibt, werde ich nun kurz die generelle Situation beschreiben, in der beide Dämonen sich frei bewegen können. Ich komme daher zu unserer allgemeinen Gleichung (5), die mit Hilfe der beiden Entropien (H und H_m) das Kriterium dafür formuliert, daß ein System sich selbst organisiert. Der Bequemlichkeit halber will ich diese Gleichung hier wiederholen und dabei gleichzeitig die Aufgaben der beiden Dämomen D_i und D_e angeben:

$$H \times \frac{\delta H_m}{\delta t} > H_m \times \frac{\delta H}{\delta t} \qquad (5)$$

H: Ergebnisse des inneren Dämons

$\frac{\delta H_m}{\delta t}$: Bemühungen des äußeren Dämons

H_m: Ergebnisse des äußeren Dämons

$\frac{\delta H}{\delta t}$: Bemühungen des inneren Dämons

Aus dieser Gleichung ist nun leicht zu ersehen, daß man den beiden Dämonen das Leben um vieles erleichtert, wenn man sie nicht zwingt, allein zu arbeiten, sondern wenn man ihnen erlaubt, zusammenzuarbeiten. Es ist erstens nicht notwendig, daß D_i die augenblickliche Entropie H stets vermindert, oder daß D_e die maximal mögliche Entropie stets vergrößert; es ist lediglich notwendig, daß das Produkt der Ergebnisse von D_i mit den Bemühungen von D_e größer ist als das Produkt der Ergebnisse von D_e mit den Bemühungen von D_i. Zweitens, wenn H oder H_m groß ist, dann kann D_e oder D_i es sich leicht machen, da seine Anstrengungen mit den entsprechenden Faktoren multipliziert werden. Dies zeigt in einer wichtigen Weise die Interdependenz dieser Dämonen. Ist nämlich D_i sehr eifrig, um ein großes H aufzubauen, dann kann D_e es sich leisten, faul zu sein, da seine Bemühungen ja mit den Ergebnissen von D_i multipliziert werden, und umgekehrt. Wenn andererseits D_e zu lange faul bleibt, hat D_i nichts, worauf er bauen kann, seine Produktion wird abnehmen und schließlich D_e zwingen, seine Tätigkeit zu verstärken, wenn das System nicht aufhören soll, ein selbst-organisierendes System zu sein.

Zusätzlich zu dieser entropischen Koppelung der beiden Dämonen findet auch eine energetische Interaktion zwischen ihnen statt, die durch die Energieanforderungen des inneren Dämons verursacht wird, der die Verschiebungen der Wahrscheinlichkeitsverteilung der Elemente des Systems bewirken soll. Dies erfordert gewisse Energie, wie wir uns vielleicht anhand des früheren Beispiels erinnern, in dem jemand die kleinen Ballons aufzublasen hatte. Da diese Energie der Umwelt entnommen worden ist, wird sie die Aktivitäten des äußeren Dämons beeinflussen, denn dieser kann in Schwierigkeiten geraten, wenn er das System mit besonders ausgesuchter Entropie zu versorgen sucht, die er aus einer energetisch erschöpften Umwelt gewinnen muß.

Zum Schluß dieser kurzen Darlegung meiner Dämonologie soll ein einfaches Diagramm die doppelte Verbindung zwischen dem inneren und dem äußeren Dämon veranschaulichen, die sie entropisch (H) und energetisch (E) interdependent macht.

Für alle jene, die diese Thematik vom Standpunkt des Physikers bearbeiten und an das Denken in den Begriffen der Thermodynamik und der statistischen Mechanik gewöhnt sind, kann unmöglich darauf verzichtet werden, auf die schöne kleine Schrift *Was ist Leben* von Erwin Schrödinger zu verweisen (1947). Alle jene unter ihnen, die dieses Buch kennen, werden sich daran erinnern, daß Schrödinger besonders zwei auszeichnende Merkmale lebender Organismen bewundert. Das eine besteht in der unglaublich hohen Ordnung der Gene, in den „erblichen Kodezeichen", wie er sie nennt, das andere in der wunderbaren Stabilität dieser organisierten Einheiten, deren empfindliche Gefüge auch dann fast gänzlich unbeeinflußt bleiben, wenn man sie thermischer Bewegung aussetzt, indem man sie – wie z.B. im Falle der Säugetiere – in einen Thermostaten eintaucht, der auf etwa 310 K gesetzt ist.

In seiner fesselnden Darlegung lenkt Schrödinger unsere Aufmerksamkeit auf zwei verschiedene Grund-„Mechanismen", durch welche geordnete Ereignisse herbeigeführt werden können: „den statistischen Mechanismus, der Ordnung aus Unordnung erzeugt, und den ... [anderen], der ‚Ordnung aus Ordnung' erzeugt".

Während der erstgenannte Mechanismus, das Prinzip „Ordnung aus Unordnung", lediglich auf „statistische Gesetze" bezugnimmt, oder, wie Schrödinger

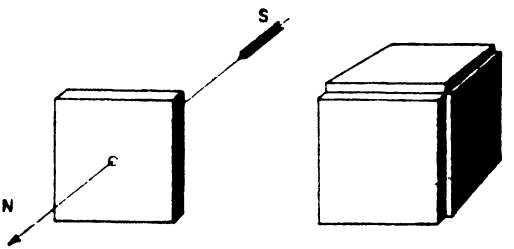
formuliert, auf „die großartige Ordnung exakter physikalischer Gesetzlichkeit, die sich aus atomarer und molekularer Unordnung ergibt", stellt der letztere Mechanismus, das Prinzip „Ordnung aus Ordnung", wiederum in seinen Worten, „den eigentlichen Schlüssel zum Verstehen des Lebens" dar. Schrödinger entwickelt dieses Prinzip bereits früh in seinem Buch sehr klar und stellt fest: „Das, wovon ein Organismus sich ernährt, ist negative Entropie". Ich glaube, daß meine Dämonen damit einverstanden wären, und ich bin es auch.

Als ich erst vor kurzem Schrödingers Bändchen wieder durchlas, fragte ich mich, wie seinen scharfen Augen entgangen sein konnte, was ich als einen „zweiten Schlüssel" zum Verstehen des Lebens oder – wenn man dies in angemessener Weise so sagen kann – selbst-organisierender Systeme betrachten würde. Auch wenn das Prinzip, das ich mir vorstelle, auf den ersten Blick als Schrödingers Prinzip „Ordnung aus Unordnung" mißverstanden werden kann, hat es tatsächlich damit überhaupt nichts gemein. Um daher den Unterschied zwischen beiden herauszuheben, möchte ich das Prinzip, das ich einführen will, als Prinzip der „Ordnung durch Störung" bezeichnen. In meinem Gasthaus ernähren sich daher selbst-organisierende Systeme nicht nur von Ordnung, für sie stehen auch Störungen auf der Speisekarte.

Ich möchte nun kurz erklären, was ich meine, wenn ich davon rede, daß ein selbst-organisierendes System sich von Störungen ernährt; ich möchte dafür ein fast triviales, aber nichtsdestoweniger amüsantes Beispiel heranziehen.

Nehmen wir an, ich besorge mir eine große Platte aus permanent magnetischem Material, das vertikal zur Oberfläche stark magnetisiert ist, und ich schneide aus dieser Platte eine große Anzahl kleiner Quadrate aus (Bild 3a). Diese kleinen Quadrate klebe ich auf die gleich großen sechs Flächen kleiner Würfel, die aus leichtem, nichtmagnetischem Material gefertigt sind (Bild 3b). Je nachdem nun, welche Seiten der Würfel den magnetischen Nordpol nach außen gerichtet aufweisen (Familie I), kann man genau zehn verschiedene Familien von Würfeln herstellen, wie Bild 4 zeigt.

Nehmen wir nun an, daß ich eine große Menge von Würfeln nehme, z.B. der Familie I, bei denen auf allen Flächen der Nordpol nach außen gerichtet ist (oder der Familie I′, bei der dies auf allen Flächen für den Südpol gilt), und sie in eine große Kiste voller kleiner Glaskiesel lege, so daß die Würfel sich unter Reibung darauf bewegen, und dann diese Kiste schüttle. Sicherlich wird nichts besonders Aufregendes geschehen: da die Würfel sich alle gegenseitig abstoßen, werden sie sich im verfügbaren Raum so verteilen, daß kein Würfel einem anderen zu nahe kommt. Wenn beim Hineinlegen der Würfel in die Kiste kein besonderes Ord-



(a) Magnetisiertes Quadrat. (b) Würfel, Familie I.

Bild 3

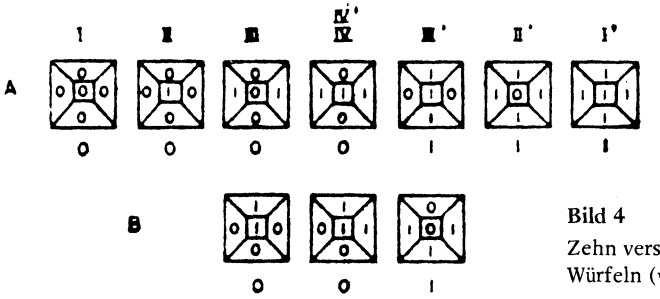


Bild 4
Zehn verschiedene Familien von Würfeln (vgl. Text)

nungsprinzip beachtet wird, bleibt die Entropie des Systems konstant oder wird im schlimmsten Fall geringfügig zunehmen.

Um dieses Spielchen noch etwas amüsanter zu machen, wollen wir annehmen, daß ich eine Menge von Würfeln zusammenstelle, die nur noch zur Hälfte zur Familie I (oder I′) gehören, zur Hälfte jedoch zur Familie II (oder II′), die dadurch charakterisiert ist, daß nur eine Fläche eine von allen anderen Flächen verschiedene, nach außen gerichtete Magnetisierung aufweist. Wird diese Menge in meine Kiste gelegt und von mir geschüttelt, werden sich alle die Würfel mit dem einen abweichenden nach außen gerichteten Pol mit überwältigender Wahrscheinlichkeit mit Mitgliedern der anderen Familie verbinden, bis fast alle Würfel gepaart sind. Da die bedingten Wahrscheinlichkeiten für ein Mitglied der Familie I, ein Mitglied der Familie II zu finden, stark zugenommen haben, hat sich die Entropie des Systems vermindert, und wir haben daher nach dem Schütteln mehr Ordnung als zuvor. Es läßt sich leicht zeigen[3)], daß der Ordnungsbetrag in unserem System sich von Null auf

$$R_\infty = \frac{1}{\log_2(en)}$$

erhöht, wenn man mit einer Populationsdichte von n Würfeln pro Inhaltseinheit beginnt.

Ich gebe Ihnen gerne zu, daß diese Zunahme an Ordnung durchaus nicht eindrucksvoll ist, besonders dann nicht, wenn die Populationsdichte hoch ist. Nun gut, wir wollen eine Population nehmen, die ausschließlich aus Mitgliedern der Familie IV B besteht, die durch entgegengesetzte Polarisierung der zwei Paare jener drei Flächen gekennzeichnet ist, die an zwei gegenüberliegenden Ecken zusammentreffen. Ich gebe diese Würfel in meine Kiste, und Sie schütteln sie. Nach einiger Zeit öffnen wir die Kiste, und statt eines Haufens von Würfeln, die irgendwie in der Schachtel übereinander liegen (Bild 5), findet sich, auch wenn Sie Ihren Augen kaum trauen, ein unglaublich geordnetes Gefüge, das sich meiner Meinung nach als durchaus geeignet erweisen könnte, in einer Ausstellung surrealistischer Kunstwerke präsentiert zu werden (Bild 6).

3) Vgl. Anhang

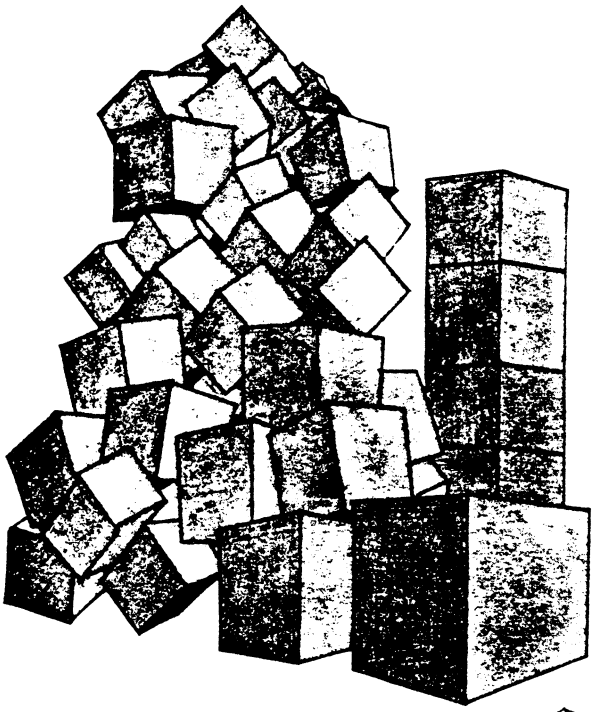


Bild 5 vorher

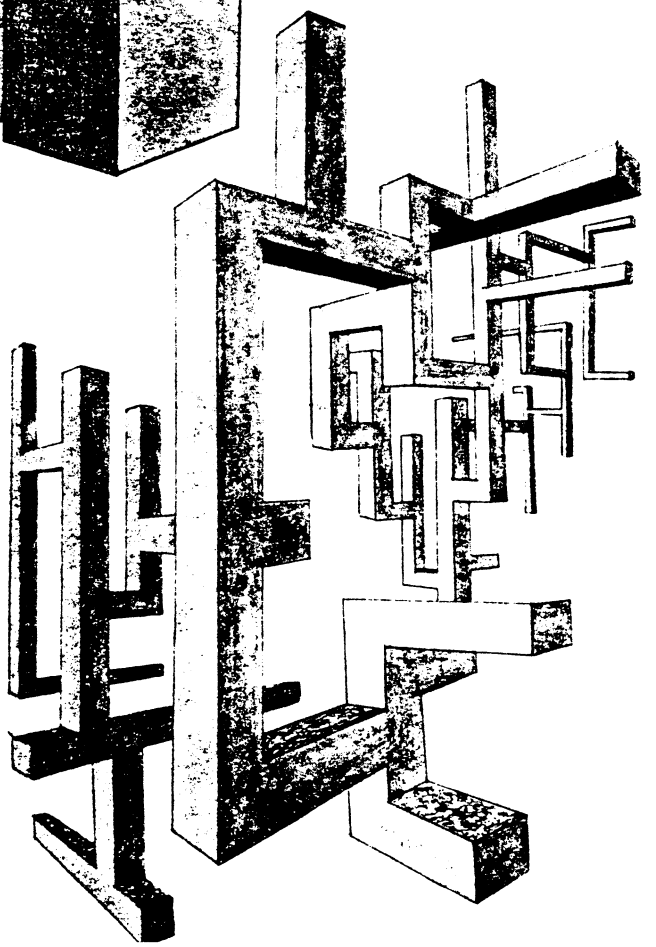


Bild 6 nachher

Hätte ich Ihnen nichts von meinem Trick mit den magnetischen Oberflächen erzählt, und würden Sie mich nun fragen, wodurch diese Würfel in diese bemerkenswerte Ordnung gebracht worden sind, dann würde ich, ohne mit der Wimper zu zucken, antworten: durch das Schütteln natürlich – und dank einiger kleiner Dämonen in der Kiste ...

Ich hoffe, mit diesem Beispiel mein Prinzip „Ordnung durch Störung" hinreichend veranschaulicht zu haben, denn diesem System wurde keine Ordnung zugegeben, lediglich billige ungerichtete Energie; dank der kleinen Dämonen in der Kiste wurden schließlich aber nur jene Störelemente ausgewählt, die zur Vergrößerung der Ordnung des Systems beitrugen. Würden wir von den Systemen der Gameten sprechen, dann wäre z.B. das Auftreten einer Mutation ein geeigneter Analogiefall.

Ich möchte daher zwei Mechanismen als wichtige Schlüssel zum Verstehen selbst-organisierender Systeme nennen: den einen können wir nach Schrödingers Vorschlag das Prinzip „Ordnung aus Ordnung" nennen, den anderen das Prinzip „Ordnung durch Störung". Beide erfordern die Zusammenarbeit unserer Dämonen, die selbst zusammen mit den Elementen unserer Systeme erzeugt werden, d.h. in einigen der immanenten Struktureigenschaften dieser Elemente gegeben sind.

Nun mag man mir vorwerfen, daß ich einen fast trivialen Fall dargestellt habe, als ich versuchte, mein Prinzip „Ordnung durch Störung" abzuleiten. Ich akzeptiere dies. Ich bin jedoch überzeugt, daß meine Position viel stärker geblieben wäre, wenn ich meinen schönen kleinen Trick mit den magnetisierten Oberflächen nicht verraten hätte. Ich bin daher den Förderern dieser Tagung sehr dankbar, daß sie Dr. Auerbach (1960) eingeladen haben, damit er über seine schönen Laborexperimente berichten kann, in denen er *in vitro* die Reorganisation von Zellen, die zuvor voneinander getrennt und gemischt worden sind, zu bestimmten festgelegten Organen demonstriert. Sollte Dr. Auerbach den Trick kennen, durch den dies bewerkstelligt wird, so hoffe ich, daß er ihn nicht verraten wird. Würde er nämlich schweigen, könnte ich meine These wieder aufnehmen, daß mein Beispiel ohne ein bestimmtes Maß an Wissen um die dabei wirksamen Mechanismen letztlich doch nicht allzu trivial war, und daß selbst-organisierende Systeme nach wie vor wundersame Dinge bleiben.

Anhang

Die Entropie eines Systems von gegebener Größe, das aus N ununterscheidbaren Elementen besteht, wird nur unter Berücksichtigung der räumlichen Verteilung seiner Elemente berechnet. Wir beginnen damit, daß wir den Raum in Z Zellen von gleicher Größe aufteilen und die Anzahl der Zellen Z_i zählen, die i Elemente enthalten (Bild 7a).

Klarerweise gilt

$$\Sigma Z_i = Z \qquad \text{(i)}$$

$$\Sigma iZ_i = N \qquad \text{(ii)}$$

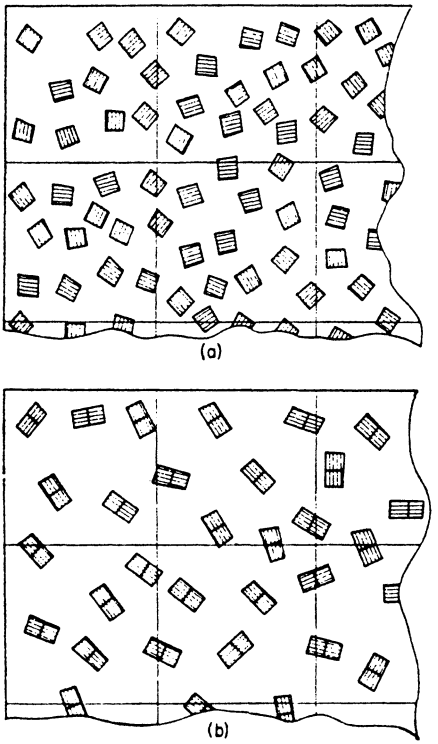


Bild 7

Die Anzahl unterscheidbarer Variationen hinsichtlich der wechselnden Anzahl von Elementen in den Zellen ist

$$P = \frac{Z!}{\Pi Z_i!} \tag{iii}$$

Daraus erhalten wir die Entropie des Systems für eine große Anzahl von Zellen und Elementen:

$$H = \ln P = Z \ln Z - \Sigma Z_i \ln Z_i. \tag{iv}$$

Im Falle der maximalen Entropie $\bar{H}$ muß gelten

$$\delta H = 0, \tag{v}$$

auch in Übereinstimmung mit den Bedingungen, die durch die Gleichung (i) und (ii) ausgedrückt werden. Wenn wir die Methode der Lagrange-Multiplikatoren anwenden, ergibt sich aus (iv) und (v) zusammen mit (i) und (ii):

$$\begin{array}{ll} \Sigma(\ln Z_i + 1)\,\delta Z_i = 0 & \\ \Sigma i \delta Z_i = 0 & \mid \beta \\ \Sigma \delta Z_i = 0 & \mid -(1 + \ln \alpha). \end{array}$$

Wenn wir hier mit den angegebenen Faktoren multiplizieren und die drei Gleichungen summieren, stellen wir fest, daß diese Summe verschwindet, wenn jeder Term in identischer Weise verschwindet. Daraus ergibt sich:

$$\ln Z_i + 1 + i\beta - 1 - \ln \alpha = 0. \tag{vi}$$

Daraus ergibt sich jene Verteilung, die H maximiert:

$$Z_i = \alpha e^{-i\beta}. \tag{vii}$$

Die zwei unbestimmten Multiplikatoren α und β können mit Hilfe der Gleichungen (i) und (ii) bewertet werden:

$$\alpha \Sigma e^{-i\beta} = Z \tag{viii}$$

$$\alpha \Sigma i e^{-i\beta} = N. \tag{ix}$$

Wenn wir uns daran erinnern, daß

$$-\frac{\delta}{\delta\beta} \Sigma e^{-i\beta} = \Sigma i e^{-i\beta},$$

erhalten wir aus (viii) und (ix) nach gewissen Umformungen:

$$\alpha = Z(1 - e^{-1/n}) \approx \frac{Z}{n} \tag{x}$$

$$\beta = \ln \left(1 + \frac{1}{n}\right) \approx \frac{1}{n}. \tag{xi}$$

Dabei ist n, die mittlere Zellpopulation oder Dichte N/Z, als groß angenommen, um diese einfachen Approximationen zu erhalten. Mit anderen Worten, es wird angenommen, daß die Zellen groß genug sind, um viele Elemente aufnehmen zu können.

Nachdem wir die Multiplikatoren α und β bestimmt haben, haben wir die wahrscheinlichste Verteilung erreicht, die entsprechend der Gleichung (vii) nunmehr folgendermaßen lautet:

$$Z_i = \frac{Z}{n} e^{-i/n}. \tag{xii}$$

Aus der Gleichung (iv) erhalten wir sodann unmittelbar die maximale Entropie:

$$\overline{H} = Z \ln(en). \tag{xiii}$$

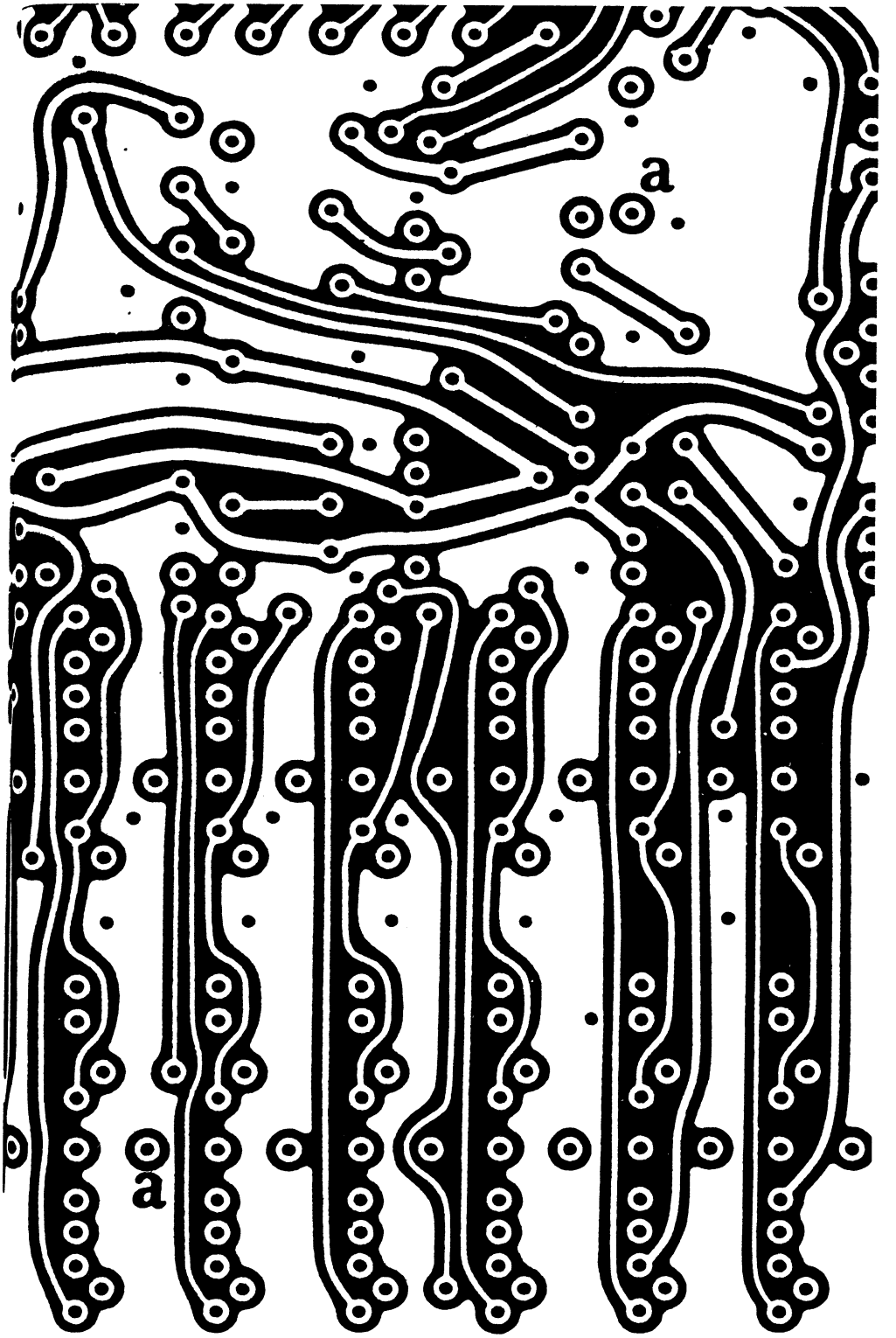
Wenn angenommen wird, daß die Elemente fähig sind, sich zu paaren (Bild 7b), ergibt sich klarerweise

$$\overline{H}' = Z \ln(en/2). \tag{xiv}$$

Wenn wir $\overline{H}$ mit H_m und $\overline{H}'$ mit H gleichsetzen, ergibt sich für den Ordnungsbetrag nach der Vereinigung

$$R = 1 - \frac{Z \ln(en)}{Z \ln(en/2)} = \frac{1}{\log_2(en)}. \tag{xv}$$

Gedächtnis ohne Aufzeichnung


a
a

Gedächtnis ohne Aufzeichnung *

H. von FOERSTER: Vielleicht sollte ich meine Position dadurch klar machen, daß ich mit einer Metapher beginne. Ich möchte bekennen, daß mir korrektes Multiplizieren große Schwierigkeiten bereitet. Ich brauche sehr lange, um zwei- oder dreistellige Zahlen zu multiplizieren, und wenn ich die gleiche Multiplikation mehrmals ausführe, bekomme ich meistens verschiedene Ergebnisse. Das ist natürlich sehr ärgerlich, und ich wollte dieses Problem daher ein für allemal erledigen, indem ich eine Aufzeichnung aller korrekten Ergebnisse anfertigte. Ich legte mir die folgende Multiplikationstabelle an: auf der linken Seite (X) bzw. oben (Y) sollten die jeweils miteinander zu multiplizierenden Zahlen, an der Schnittstelle der entsprechenden Zeilen und Spalten sollte ihr Produkt (XY) eingetragen werden (Tab. 1).

Tabelle 1

X · Y		0	1	2	3	.	.	.	.	.	Y
X	0	0	0	0	0	.	.	.			
	1	0	1	2	3	.	.	.			
	2	0	2	4	6	.	.	.			
	3	0	3	6	9	.	.	.			
	.	.	.	.	.	.	.	.			
	.	.	.	.	.	.	.	.			

Bei der Anfertigung dieser Tabelle stellte sich die Frage, wieviel Papier ich brauchen würde, um alle Faktoren X und Y bis zu einer Größenordnung von, sagen wir, n Dezimalziffern unterzubringen. Tippt man die Zahlen in normaler Größe auf die üblichen Blätter 21 X 27 cm, dann läßt sich die Dicke D des Buches, das meine Multiplikationstabelle für die Zahlen mit bis zu n Dezimalziffern enthält, näherungsweise berechnen wie folgt:

$$D = n \cdot 10^{2n-6} \text{ cm.}$$

Eine 100 X 100 Multiplikationstabelle ($100 = 10^2$; $n = 2$) füllt z.B. ein „Buch" von der Dicke

$$D = 2 \cdot 10^{4-6} = 2 \cdot 10^{-2} = 0{,}02 \text{ cm} = 0{,}2 \text{ mm.}$$

Mit anderen Worten, diese Tabelle kann auf ein einziges Blatt Papier gedruckt werden.

---

* Bearbeitete Nachschrift eines Vortrags vom 2. Oktober 1963 auf der "First Conference on Learning, Remembering, and Forgetting", Princeton/New Jersey.

PRIBRAM:

Aber Sie sagten doch, daß Sie nicht multiplizieren können!

von FOERSTER:

Das ist richtig. Ich arbeite daher nur mit den Exponenten, und dafür brauche ich bloß zu addieren.

Ich schlage nun vor, meine Tabelle auf Multiplikationen von zehnstelligen Zahlen zu erweitern. Dies ist ein sehr bescheidener Wunsch; eine solche Tabelle kann schon nützlich sein, wenn man seine Steuererklärung macht. Mit unserer Formel für D erhalten wir für n = 10:

$$D = 10 \cdot 10^{20-6} = 10^{15} \text{ cm}.$$

Mit anderen Worten, diese Multiplikationstabelle muß auf einem Bücherregal untergebracht werden, das 10^{15} cm lang ist, d.i. etwa 100mal die Entfernung zwischen Sonne und Erde oder die Länge etwa eines Lichttages. Ein Bibliothekar, der sich mit Lichtgeschwindigkeit bewegt, braucht im Durchschnitt einen halben Tag, um eine einzige Eintragung in dieser Tabelle nachzusehen.

Dies schien mir nun kein sehr praktisches Verfahren zu sein, meine Information, d.h. die Ergebnisse aller Multiplikationen zehnstelliger Zahlen, aufzubewahren. Da ich diese Information aber ganz dringend benötigte, mußte ich mich nach einer anderen Methode umsehen. Ich kam dabei auf ein Gerät, das 12 X 12 X 19 cm groß ist und 20 kleine Räder enthält, denen allen die Ziffern von 0 bis 9 aufgeprägt sind. Diese Räder sitzen auf einer Achse und sind durch Zähne und Stifte auf ausgeklügelte Art so miteinander verbunden, daß man nach entsprechendem Drehen einer Kurbel das gewünschte Ergebnis einer Multiplikation durch ein Fenster von den Rädern selbst ablesen kann.

Das ganze Gerät ist äußerst billig und benötigt im Durchschnitt nur 50 Kurbeldrehungen, um jedes gewünschte Ergebnis einer Multiplikation von zwei zehnstelligen Zahlen zu erzielen.

Die Antwort auf die Frage, ob ich die Information einer $10^{10} \times 10^{10}$ Multiplikationstabelle in der Form eines 21 X 27 cm Buches mit einer Dicke von ca. zehn Milliarden km „speichern" sollte oder in der Form eines kleinen handbetriebenen Tischrechners, ist daher, so meine ich, ganz eindeutig. Nun läßt sich aber einwenden, daß der Rechner diese Information nicht „speichert", sondern jede Aufgabe durch eine begrenzte Menge von Operationen einzeln löst. Wenn ich die Kurbel drehe, dann gebe ich dem Rechner nichts weiter als die „Adresse" des Ergebnisses, welches ich sodann unmittelbar entnehmen kann — ohne daß der „Rechner" irgend etwas tut —, indem ich die zum Schluß erreichte Stellung der Räder ablese. Wenn ich auf diese Weise Information entnehmen kann, dann muß sie doch zuvor in das System hineingegeben worden sein. Aber wie? Ganz offensichtlich wird die Information im Rechner in struktureller Weise gespeichert. Durch die Art, in der die Räder interagieren, durch Einkerbungen und Verbindungsstifte, wird die gesamte Information, die für das Erreichen des richtigen Ergebnisses notwendig ist, im Konstruktionskode der Maschine, oder, um es biologisch zu sagen, in ihrem genetischen Kode, niedergelegt.

Wenn man mich nun bittet, ein „Gehirn" zu bauen, das zu ähnlichen oder noch komplizierteren Kunststücken fähig ist, dann würde ich eher an ein kleines

und kompaktes Rechengerät denken als an Tabellierungsmethoden, die allzu schnell jeden vernünftigen Rahmen überschreiten.

Ich habe im Verlauf dieser Tagung den Eindruck gewonnen, daß Sie, meine Herren, in zahlreichen Beispielen und Aussagen beträchtliches Beweismaterial dafür präsentiert haben, daß das Nervensystem wie ein Rechner operiert. Zu meinem großen Erstaunen jedoch schienen Sie mir – soweit ich einige der Diskussionspunkte überhaupt verstehen konnte – oft so zu argumentieren, als ob das Gehirn ein Speicher für eine gigantische Tabelle wäre. Und um bei meinem Bild zu bleiben: Ihre Überlegungen scheinen sich ausschließlich damit beschäftigt zu haben, ob die Symbole in meiner Multiplikationstabelle in grüner oder roter Farbe oder vielleicht in Blindenschrift gedruckt sind, statt daß Sie sich gefragt hätten, ob der Zifferntransport in meinem Tischrechner durch ein Friktions- oder ein Zahnradgetriebe bewerkstelligt wird.

Ich muß zugeben, daß meine Metapher bis jetzt immer noch äußerst unzureichend ist, denn mein Rechner ist eine deterministische und starre Angelegenheit und seine Verhaltensregeln sind alle *a priori* festgelegt. Ein solches System kann durch Erfahrung nichts lernen und sollte daher in einer Tagung zum Thema „Gedächtnis" überhaupt nicht erwähnt werden.

Ich werde daher meine Metapher etwas erweitern und vorschlagen, einen Rechner zu bauen, der die Operationen, deren Beherrschung ich von ihm fordere, durch eigene Erfahrung erst zu lernen hat. Mit anderen Worten, ich habe mir das Problem gestellt, einen adaptiven Rechner zu bauen. Erlauben Sie mir jedoch einige Vorbemerkungen, bevor ich versuche, eine Lösung dieses Problems zu entwickeln.

Zunächst möchte ich nochmal von der Versuchung sprechen, vergangene Erfahrung als eine Art Aufzeichnung anzusehen. Diese Sehweise bietet sich natürlich an, da ein kumulativer Speicher sehr leicht zu bauen ist. Man zeichnet einfach immerfort auf ... Dabei vernachlässigt man aber gewöhnlich alle jene Probleme, die sich ergeben, wenn man versucht, solche Aufzeichnungen für irgendwelche Zwecke zu *nutzen*! Doch lassen wir diesen kitzligen Punkt einmal beiseite! Fragen wir vielleicht, *wie* man aufzeichnen soll. Wiederum kann man überlegen, um bei meiner Metapher zu bleiben, ob man Tinte benützen sollte, die nach gewisser Zeit verbleicht, wenn sie nicht verstärkt wird, ob gültige oder ungültige Eintragungen durch Plus- bzw. Minuszeichen gekennzeichnet werden sollen, oder ob die Wichtigkeit einer Eintragung durch Fettdruck ausgedrückt werden sollte usw. Fragen dieser Art ergeben sich, wenn wir uns mit der technischen Herstellung einer großen Aufzeichnung befassen. Die Probleme sind jedoch völlig anderer Art, wenn wir uns dem Bau eines adaptiven Rechengerätes zuwenden, dessen innere Struktur durch seine Interaktionen mit einer Umwelt modifiziert wird.

Ich glaube, daß viele der Ausführungen auf dieser Tagung sich mit dem Problem befaßt haben, *wie* die Aufzeichnung zu schreiben sei, und nicht mit dem Problem, wie die Struktur eines Rechners so angelegt werden kann, daß seine Arbeitsweise sich mit zunehmender Erfahrung verändert. Es war für mich jedoch interessant zu beobachten, daß diese Ausführungen gewöhnlich dann gemacht wurden, wenn die Redner sich überlegten, wie das System arbeiten „sollte", und nicht dann, wenn sie sich mit dem tatsächlichen Operieren des Systems befaßten. Dies war z.B. der Fall, als ich von Sir John Eccles zu hören meinte, daß Lernen

eine Verstärkung der synaptischen Leistung durch ihren Gebrauch *erfordert*. Wie er jedoch mit seinem interessanten Beispiel der Reflexaktion eines Muskels, der eine gewisse Zeit vom Knochen abgetrennt war, zeigte, arbeiten diese elenden Zellen gerade auf entgegengesetzte Weise: ihre Leistung nahm mit der Dauer des Ruhezustandes zu. Oder, obwohl Dr. Krüger, wenn ich mich richtig erinnere, festgestellt hat, daß Vergessen die Degeneration von Neuronen *erfordert*, entnehme ich doch seinen Ausführungen, daß es sehr schwer zu sein scheint, diese Zellen überhaupt zum Sterben zu bringen.

Es ist völlig klar, daß die Ausführungen über die Art dieser Bestandteile von der Vorstellung bestimmt sind, daß sie in einem adaptiven Aufzeichnungsgerät verwendet werden sollen. Ich schlage nun vor, kurz zu überlegen, ob die Art und Weise, in der sich diese Bestandteile tatsächlich verhalten, genau der Art und Weise entspricht, in der sie sich verhalten müßten, wenn wir sie als Bausteine für ein adaptives Rechengerät verwendeten. So kann z.B. die Degeneration von Neuronen – wenn es sie gibt – ein wichtiger Mechanismus der Erleichterung des Lernens sein, wenn Lernen mit der Unterdrückung irrelevanter Reaktionen zusammenhängt, wie Sir John dies vorhin festgestellt hat. Andererseits kann die Verstärkung einer Verbindung, wie sie durch eine verlängerte Ruheperiode bewirkt wird, in Verbindung mit bestimmten inhibitorischen Aktionen als ein Mechanismus des „Vergessens" benutzt werden.

McCONNELL:

Ich möchte ungern ihre hübsche Analogie kaputt machen, aber gibt es nicht viele Fälle, in denen eine Aufzeichnung einfach viel effektiver wäre? Ich denke etwa an Primzahlen?

von FOERSTER:

Richtig. Aber bitte erlauben Sie mir, meine Geschichte ein wenig weiterzuspinnen. Ich werde gleich die einschränkenden Bedingungen für meinen Rechner erheblich verschärfen. Auf Ihre Bemerkung komme ich später zurück. Verzeihen Sie mir, daß ich meine Metapher so langsam entwickele.

McCONNELL:

Wir verfangen uns bereits in Ihren Fäden.

von FOERSTER:

Das sollten Sie nicht! Haben Sie noch ein bißchen Geduld, dann können Sie über mich herfallen. Ich gebe Ihnen sicher eine Reihe von Gelegenheiten.

Meine zweite Vorbemerkung mit Bezug auf die Konstruktion eines adaptiven Rechners betrifft die Auswahl einer guten Strategie für die Bearbeitung dieses Problems. Glücklicherweise hat uns Sir John in der Kurzfassung, die vor der Tagung verteilt wurde, eine ausgezeichnete Leitlinie gesetzt: „Lernen bedeutet Selektivität der Reaktion; es ist zu vermuten, daß jede Inhibition in besonderem Maße die Unterdrückung irrelevanter Reaktionen bewirkt". Mit anderen Worten, Lernen bedeutet selektives Operieren; um jedoch die Ergebnisse solcher Operationen zu erzielen, bedarf es einer Rechenvorrichtung, um derartige selektive Operationen durchzuführen.

PRIBRAM:

Sir John, war Ihnen klar, was Sie sagten?

ECCLES:

Nein. (Lachen)

von FOERSTER:

Wie würden Sie das andernfalls machen? Wie würden Sie selektieren? Wie kämen selektive Manipulationen zustande?

ECCLES:

Wie macht das ein Tier?

von FOERSTER:

Ich glaube, daß im Laufe dieser Tagung eine eindrucksvolle Reihe von Vorschlägen dazu gemacht worden ist. Ich denke z.B. an Ihre Demonstration der Veränderungen der Effikanz synaptischer Verbindungen aufgrund unterschiedlicher Stimuli, Sir John. Statt diese Veränderungen als Speicherpunkte für bestimmte vorübergehende Ereignisse anzusehen, schlage ich vor, sie als Veränderungen der Transferfunktionen eines Rechenelements zu interpretieren. Mit anderen Worten, ich schlage vor, diese lokalen Veränderungen als – zugegeben winzige – Modifikationen der Reaktionsmerkmale des Systems als ganzen aufzufassen. Dr. Uttley hat in seinem Vortrag jede Verbindung als Rechenelement aufgefaßt, das mit bedingter Wahrscheinlichkeit funktioniert. Dr. Hydéns höchst kompliziertes Modell besteht darin, die passenden Proteine mit Hilfe kodierter DNS- und RNS-Matrizen zu berechnen. Der Biochemiker würde wahrscheinlich Begriffe wie „formen" oder „synthetisieren" statt „rechnen" verwenden, diese Ausdrücke sind jedoch in einem abstrakten Sinn äquivalent.

Lassen Sie mich nun auf das ursprüngliche von mir gestellte Problem zurückkommen, nämlich auf die Konstruktion einer Rechenvorrichtung, die ihre eigene innere Organisation aufgrund von Interaktionen mit ihrer Umwelt verändert. Ich glaube, es ist ganz klar, daß ich zwei Fragen eliminieren muß, um mit meiner Konstruktionsarbeit überhaupt voranzukommen: 1. An welche Art von Umwelt ist mein Rechner gekoppelt? 2. Was soll mein Rechner von dieser Umwelt lernen?

Solange ich in der Lage bin, die Regeln festzulegen, die die Ereignisse in dieser Umwelt und die von meinem Rechner zu bewältigenden Aufgaben bestimmen, dürfte es nicht allzu schwierig sein, das entsprechend angemessene System zu entwerfen. Nehmen wir für den Augenblick einmal an, daß die Umwelt ganz simpel so gestaltet ist, daß sie meinen Rechner immer dann mit einem merklichen Energiebetrag belohnt, wenn dieser das richtige Ergebnis einer Multiplikationsaufgabe liefert, die ihm von der Umwelt gestellt worden ist. Natürlich könnte ich meinen alten Tischrechner unmittelbar an diese Umwelt anschließen, gäbe es nicht einen Haken: das Zahlensystem, in dem die Umwelt ihre Fragen formuliert, ist nicht *a priori* festgelegt. Es kann ein Dezimalsystem sein, ein Binärsystem, eine Abbildung von Produkten von Primzahlen, oder – wenn wir besonders häßlich sein wollen – die Umwelt kann ihre Probleme in römischen Ziffern stellen. Auch wenn ich annehme, daß mein Rechner die platonische Idee der Multiplikation eingebaut hat, muß er das Zahlensystem lernen, um in dieser Umwelt erfolgreich zu sein.

Damit ist meine Aufgabe nun hinreichend genau bestimmt. Ich kenne die Struktur der Umwelt, ich weiß, was mein System zu lernen hat, und ich kann nun darüber nachzudenken beginnen, wie dieses Problem zu lösen ist.

Ohne daß wir uns nun weiter damit amüsieren, wie wir das Problem lösen, dieses profane Maschinchen zu bauen, lassen Sie mich gleich auf das eigentliche Problem kommen, dem wir uns gegenübersehen, auf das Problem nämlich, wie lebende Organismen es schaffen, sich selbst in einer Umwelt am Leben zu erhalten, die alles eher als simpel ist.

Die Frage nach der Umwelt, an die unsere Systeme gekoppelt sind, ist nun insofern beantwortet, als es sich dabei um die Natur mit all ihren Unvorhersagbarkeiten, aber auch mit den zwingenden Gesetzmäßigkeiten handelt, die in den Gesetzen der Physik oder der Chemie niedergelegt sind.

Damit sind wir in der Lage, nun die zweite Frage zu stellen: Was müssen unsere Organismen lernen können? Vielleicht kann diese Frage eher beantwortet werden, wenn wir uns zuerst fragen: „Warum sollen diese Organismen überhaupt etwas lernen?" Ich glaube, daß wir zum harten Kern des Problems vorgestoßen sind, das uns alle hier zusammengeführt hat, wenn wir eine ausreichende Antwort auf diese Frage finden. Mit meinem Vorschlag einer Antwort auf diese Frage bin ich zum zentralen Punkt meines Vortrags gekommen. Ich glaube, daß der eigentliche Grund dafür, daß diese Systeme überhaupt etwas lernen müssen, darin besteht, daß solches Lernen sie instandsetzt, induktive Schlüsse zu ziehen. Mit anderen Worten, das System muß zur Verbesserung seiner Überlebenschancen imstande sein, zukünftige Ereignisse aus vergangener Erfahrung zu berechnen, es muß ein „(Er-)Rechner induktiver Schlüsse" sein. Darüber hinaus liegt auf der Hand, daß nur ein System mit einem Gedächtnis imstande ist, solche induktiven Schlüsse zu ziehen, denn es ist unmöglich, aus dem Zeitstück gegenwärtiger Ereignisse allein Schlüsse über nachfolgende Ereignisse abzuleiten, wenn frühere Zustände der Umwelt nicht in Betracht gezogen werden können.

Ich habe damit die nähere Bestimmung meiner Aufgabe abgeschlossen: Ich möchte einen (Er-)Rechner induktiver Schlüsse bauen, dessen Zunahme an innerer Organisation Ungewißheiten hinsichtlich der Vorhersagen zukünftiger Ereignisse in seiner Umwelt beseitigt.

An diesem Punkt möchte ich an die Situation erinnern, in der wir ein Rechengerät noch mit einem Aufzeichnungsgerät gekoppelt haben, um unser Gedächtnisproblem zu bewältigen. Aus der eben beschriebenen Aufgabe ergibt sich jedoch klar, daß eine Aufzeichnung der Vergangenheit, wie detailliert und dauerhaft sie auch immer sein möge, völlig wertlos ist. Sie ist tot. Sie gibt uns nicht den geringsten Hinweis auf zukünftige Ereignisse, wenn wir nicht einen Dämon einsetzen, der ständig über diese Aufzeichnungen hinwegfegt, mit Lichtgeschwindigkeit für jede Eintragung einen bestimmten Wert berechnet, diese Werte in einer Menge von Selektionsoperationen miteinander vergleicht und schließlich daraus die Wahrscheinlichkeitsverteilung der nächsten zukünftigen Ereignisse berechnet. Dieser Dämon muß dies alles außerdem *zwischen* den zeitlichen Augenblicken erledigen! Wenn wir weiter darauf bestehen, Aufzeichnungen zu machen, verlagern wir lediglich unser Problem des Gedächtnisses auf die Fähigkeiten eines solchen Dämons, der nun seinerseits die Aufgabe hat, als ein (Er-)Rechner induktiver Schlüsse zu arbeiten. Ich kann daher die Aufzeichnungen ohne weiteres wegwerfen und über die Konstruktion dieses Dämons nachdenken, der ja gar nicht auf die Aufzeichnung der Ereignisse zu schauen braucht, sondern nur auf die Ereignisse selbst.

Damit habe ich nun die metaphorische Phase meines Vortrags beendet, die, wie ich hoffe, in qualitativer Weise meine Position umschrieben und mein Problem geklärt hat. Ich möchte nun dazu übergehen, das Problem aus quantitativer Perspektive zu behandeln. Mit anderen Worten, bevor ich mich an die tatsächliche Konstruktion eines solchen (Er-)Rechners induktiver Schlüsse mache, wird es sich empfehlen, auf die eine oder andere Weise abzuschätzen, *wie viel* an innerer Organisation unser Rechner im Laufe seiner Interaktionen mit der Umwelt aufbauen soll und *wie viel* an Ungewißheit mit Bezug auf zukünftige Ereignisse er durch diesen Erwerb höherer Ordnungszustände beseitigen können soll.

Es ist glücklicherweise möglich, zwei entscheidende Begriffe meines Argumentationsgangs in präziser quantitativer Weise zu definieren. Dies ist einmal der Begriff der Ungewißheit, zum anderen der Begriff der Ordnung. In beiden Fällen ist es möglich, adäquate Maßfunktionen zu definieren, die die Übersetzung meines Problems in einen mathematischen Formalismus gestatten. Da die ganze dafür benötigte mathematische Maschinerie in der bereits bekannten „Informationstheorie" vollständig entwickelt ist, mag es an dieser Stelle ausreichen, auf die wichtigste einschlägige Literatur zu verweisen (Shannon/Weaver 1949; Brillouin 1962; Defares/Sneddon 1961), worunter meiner Meinung nach die Darstellung des verstorbenen Henry Quastler (1958) für den biologisch Interessierten am attraktivsten ist. Mit der Erlaubnis unseres Vorsitzenden möchte ich einige der Grundbegriffe dieser Theorie erneut definieren, um all jenen entgegenzukommen, die ihr Gedächtnis aufgefrischt haben möchten, ohne andere Quellen konsultieren zu müssen.

Lassen Sie mich rasch einige der Grundbegriffe der Informationstheorie vorstellen. Gleichzeitig möchte ich alle jene um Verständnis bitten, die in dem nun folgenden Schatten einer Skizze dieser Theorie die notwendige Strenge vermissen.

Der grundlegendste Schritt in einer mathematischen Theorie der Information ist die Entwicklung eines Maßes für den Unsicherheitsbetrag einer Situation als ganzer, oder – wie ich es formulieren will – für die Unsicherheit eines „wohldefinierten Universums". Die Definition dieses hier in Frage stehenden Universums kann auf verschiedenen Ebenen vorgenommen werden. Der erste Schritt in seiner Definition besteht darin, für dieses Universum eine endliche Zahl unterscheidbarer Zustände zu bestimmen, die zudem alle die Zustände umfassen, die das Universum einnehmen kann. Altgediente, aber anschauliche Beispiele für ein derartiges Universum sind etwa ein Würfel mit seinen sechs Flächen oder eine Münze mit ihren zwei Seiten. Die Seite der Münze bzw. die Würfelfläche, die nach einem Wurf oben liegt, repräsentiert einen unterscheidbaren Zustand in diesen „Universen". Aufgrund der Unterscheidbarkeit der einzelnen Zustände ist es möglich, diese jeweils zu benennen, etwa als S_1, S_2, S_3 usw. Allgemein können wir einen Zustand mit S_i benennen, wobei i alle ganzen Zahlen von 1 bis n durchlaufen kann, wenn unser Universum durch präzise n Zustände definiert ist. So gilt etwa für eine Münze:

$$n = 2; \quad \begin{matrix} S_1 = \text{Kopf} \\ S_2 = \text{Zahl.} \end{matrix}$$

Ähnliches gilt für den Würfel mit n = 6, bei dem die Namen der Zustände S_i – der Einfachheit halber – mit der Anzahl der Augen i übereinstimmen, die der Würfel jeweils zeigt.

Solange wir es nicht mit einem völlig deterministischen Universum zu tun haben, d.h. mit einem Universum, in dem für jeden einzelnen Zustand ein und nur ein möglicher Nachfolgezustand existiert (und unsere gegebenen Beispiele sind gewiß nicht von dieser Art), verbinden wir mit einem solchen indeterministischen Universum intuitiv einen gewissen Betrag an Unsicherheit. Wir können etwa mit Bezug auf die Münze sagen, daß wir das Ergebnis eines einzelnen Wurfes nicht vorhersagen können, wir sind dabei aber viel weniger unsicher hinsichtlich der Gesamtsituation, wenn wir die Münze mit dem Würfel vergleichen, der eine viel größere Mannigfaltigkeit möglicher Ergebnisse zeigt. Die Frage, *wie viel* an Unsicherheit mit diesen unterschiedlichen Situationen verbunden werden kann, führt zum zweiten Schritt in der Definition unserer Universen. Da Wahrscheinlichkeitsüberlegungen klarerweise diese Unsicherheiten bestimmen, schlage ich vor, jedem Zustand S_i in unserem Universum die Wahrscheinlichkeit p_i seines Auftretens zuzuordnen. Da unser Universum präzise n Zustände aufweist, und da es daher in jedem Zeitpunkt in einem der Zustände sein muß, haben wir natürlich die Gewißheit, daß es sich in einem gegebenen Zeitpunkt tatsächlich in einem bestimmten seiner Zustände befindet:

$$p_1 + p_2 + \ldots + p_n = \sum_1^n p_i = 1. \qquad (1)$$

In der simplen Situation eines Universums, in dem alle Wahrscheinlichkeiten p_i gleich sind, etwa p_u – wie dies für eine „ehrliche" Münze oder einen „ehrlichen" Würfel der Fall ist –, lautet die obige Gleichung einfach

$$np_i = 1,$$

und die Wahrscheinlichkeit für einen einzelnen Zustand S_i ist exakt die Umkehrung der Anzahl der Zustände:

$$p_u = p_i = \frac{1}{n},$$

oder mit Bezug auf unsere zwei Beispiele:

$$p_{\text{Münze}} = \frac{1}{2} \quad \text{und} \quad p_{\text{Würfel}} = \frac{1}{6}.$$

Wenn wir nun jedem Universum ein Maß der Unsicherheit zuordnen wollen, dann erscheint es zumindest plausibel, daß dieses Maß die Wahrscheinlichkeiten, bzw. die Ungewißheiten, wenn Sie wollen, aller dieses Universum definierenden Zustände berücksichtigen muß. Mit anderen Worten, daß Maß der Unsicherheit – gewöhnlich mit H bezeichnet – eines bestimmten Universums sollte eine Funktion aller p_i sein:

$$H = H(p_1; p_2; p_3; \ldots p_i \ldots p_n). \qquad (2)$$

Da es unendlich viele Funktionen gibt, aus denen man eine auswählen kann, wie z.B.

$$\sum_1^n p_i^2; \quad \sum_1^n e^{p_i}; \quad \prod_1^n \log p_i; \quad \text{etc., etc., } \ldots$$

sind wir nun in der Lage, bestimmte Bedingungen einzuführen, die wir aus intuitiven Gründen bei der endgültigen Entscheidung für ein Maß der Unsicherheit eines Universums erfüllt sehen möchten. Es kann nicht genug betont werden, daß die Wahl dieser Bedingungen mehr oder minder beliebig ist, ihre Rechtfertigung liegt ausschließlich in ihren Konsequenzen.

Eine dieser Bedingungen kann vernünftigerweise darin bestehen, daß das Maß dieser Unsicherheiten in gewissem Sinne unsere Intuition des Betrags dieser Unsicherheiten spiegeln sollte. Mit anderen Worten, größere Unsicherheit sollte durch ein höheres Maß an Unsicherheit repräsentiert werden.

Eine weitere Bedingung kann sein, daß das Unsicherheitsmaß in einem deterministischen Universum verschwinden muß (H = 0), d.h. in einem Universum, in dem es keine Unsicherheiten gibt.

Wir können schließlich vorschlagen, daß das Unsicherheitsmaß für zwei voneinander unabhängige Universen, U_1 und U_2, die Summe der Unsicherheitsmaße jedes einzelnen Universums sein soll. Mathematisch wäre dies:

$$(H)\ U_1\ \&\ (U_2) = H(U_1) + H(U_2). \tag{3}$$

Aus dieser letzten Bedingung können wir auch einen Hinweis auf die Form unserer Maßfunktion ableiten. Erinnern wir uns einen Augenblick an unsere zwei Beispiele, die Münze und den Würfel. Wir fordern, daß das Maß für die Unsicherheit eines Universums, das aus einer Münze und einem Würfel besteht, gleich ist der Summe der Maße, die die Unsicherheiten des Münzenuniversums bzw. des Würfeluniversums ausdrücken. Da im Verbunduniversum die Anzahl der Zustände n_{CD} das Produkt der Anzahl der Zustände der Teiluniversen n_C, n_D ist:

$$n_{CD} = n_C \cdot n_D = 2 \cdot 6 = 12;$$

und da alle Wahrscheinlichkeiten gleich sind:

$$p_{CD} = \frac{1}{n_{CD}} = \frac{1}{n_C \cdot n_D} = p_C \cdot p_D,$$

ergibt sich mit Hilfe eines postulierten Additionstheorems unabhängiger Universen:

$$H(p_{CD}) = H(p_C \cdot p_D) = H(p_C) + H(p_D).$$

Diese Beziehung drückt aus, daß die gewünschte Maßfunktion, H, als Produkt zweier Faktoren gleich sein muß der Summe der Maßfunktionen der Faktoren. Es gibt nur eine mathematische Funktion, die dieser Bedingung genügt, nämlich die logarithmische Funktion. Wir schreiben daher versuchsweise an:

$$H(p) = K \cdot \log(p),$$

wobei k eine noch undefinierte Konstante darstellt. Wir verifizieren nun die obige Relation $H(p_{CD})$:

$$\begin{aligned} H(p_{CD}) &= k \cdot \log(p_{CD}) = k \cdot \log(p_C \cdot p_D) \\ &= k \cdot \log(p_C) + k \cdot \log(p_D) \\ &= H(p_C) + H(p_D). \end{aligned}$$

Q.E.D.

Da wir andererseits intuitiv glauben, daß ein Universum, das mehr an Zuständen einnehmen kann als ein anderes, auch mit einem höheren Grad an Unsicherheit ausgestattet ist, sind wir gezwungen, unsere noch undefinierte Konstante k mit einem negativen Vorzeichen zu versehen:

$$H(p) = -k \cdot \log(p) = k \cdot \log\left(\frac{1}{p}\right) = k \cdot \log(n).$$

In der einfachen Situation von Universen, deren Zustände gleichwahrscheinlich sind, haben wir damit den Schluß erreicht, daß eine adäquate Funktion zur Berechnung eines Unsicherheitsmaßes von logarithmischer Form ist:

$$H(p) = k \cdot \log(n) = -k \cdot \log(p). \tag{4}$$

Nun sind Situationen, in denen alle Zustände gleichwahrscheinlich sind, ziemlich rar, und wir müssen uns mit dem allgemeinen Fall beschäftigen, in dem jeder Zustand S_i mit einer Wahrscheinlichkeit p_i versehen ist, die der Wahrscheinlichkeit anderer Zustände gleich oder auch nicht gleich sein kann. Da sich das Unsicherheitsmaß in der Situation der Gleichwahrscheinlichkeit als proportional dem Logarithmus des reziproken Wertes der Wahrscheinlichkeit eines einzelnen Zustandes erwiesen hat, wird die Annahme nahegelegt, daß das Unsicherheitsmaß im allgemeinen Fall dem *Mittelwert* des Logarithmus des reziproken Wertes jedes einzelnen Zustandes proportional ist,

$$H = -k \cdot \overline{\log p_i}. \tag{5}$$

(Der Balken zeigt den Mittelwert von $\log(p_i)$ für alle Zustände n an.)

Die Berechnung eines Mittelwertes ist einfach. Man nehme N Stöcke, die aus n Gruppen bestehen, von denen jede N_i Stöcke von der Länge l_i aufweist. Wie groß ist der Mittelwert ihrer Länge? Klarerweise: die Gesamtlänge aller Stöcke dividiert durch die Anzahl der einzelnen Stöcke:

$$\bar{l} = \frac{\sum_{1}^{n} N_i \cdot l_i}{N}.$$

Wir nennen nun p_i die Wahrscheinlichkeit des Auftretens eines Stockes mit der Länge l_i:

$$p_i = N_i/N,$$

und der Ausdruck für die mittlere Länge wird zu

$$\bar{l} = \sum_{1}^{n} \frac{N_i}{N} \cdot l_i = \sum_{1}^{n} p_i \cdot l_i.$$

Mit anderen Worten, der Mittelwert einer Menge von Werten wird einfach durch die Summe der Produkte der verschiedenen Werte mit der Wahrscheinlichkeit ihres Auftretens berechnet. Der Mittelwert der verschiedenen Werte von $\log p_i$ ist daher einfach

$$\overline{\log p_i} = \sum_{1}^{n} p_i \log p_i,$$

und das Unsicherheitsmaß wird zu

$$H = -k\,\overline{\log p_i} = -k \sum_1^n p_i \log p_i.$$

Diese Maßfunktion erfüllt alle die von uns geforderten Bedingungen. Sie reduziert sich auf die einfache Gleichung für gleichwahrscheinliche Zustände, denn mit

$$p_i = \frac{1}{n},$$

ergibt sich

$$H = -k \sum_1^n \frac{1}{n} \log \frac{1}{n} = -k \frac{n}{n} \log \frac{1}{n} = k \ln n.$$

Daraus folgt, daß H mit einer zunehmenden Zahl von gleichwahrscheinlichen Zuständen in monotoner Weise selbst zunimmt. Außerdem verschwindet H, wenn es Sicherheit gibt. Dies werden wir dadurch ausdrücken, daß wir annehmen, daß es nur einen Zustand gibt, der mit Sicherheit auftritt, d.h. $p_i = 1$, während alle anderen die Wahrscheinlichkeit 0 haben. Daraus folgt

$$H = -k \left[1 \cdot \log 1 + \sum_2^n 0 \cdot \log 0 \right] = 0.$$

weil

$$1 \log 1 = 0.$$
$$0 \log 0 = 0.$$

Da der letzte Ausdruck nicht unmittelbar einleuchtend ist, da $\log 0 = -\infty$, wollen wir rasch mit Hilfe von l'Hospital's Regel zeigen, daß

$$\lim_{x \to 0} (x \cdot \log x) = 0;$$

$$\lim_{x \to 0} (x \cdot \log x) = \lim \frac{\log x}{\left(\frac{1}{x}\right)} = \lim \frac{\frac{d \log x}{dx}}{\frac{d}{dx}\left(\frac{1}{x}\right)}$$

$$= \lim \frac{1/x}{-1/x^2} = \lim_{x \to 0} (-x) = 0. \qquad \text{Q.E.D.}$$

Schließlich zeigen wir, daß auch das Additionstheorem für voneinander unabhängige Universen aufrechterhalten wird. Seien p_i und p_j die Wahrscheinlichkeiten der Zustände der beiden Universen U_1 und U_2, dann sind die Wahrscheinlichkeiten des aus beiden kombinierten Universums

$$p_{ij} = p_i \cdot p_j.$$

Seien n und m die Anzahl der Zustände, die U_1 und U_2 entsprechen. Für U_1 und U_2 ist die Anzahl der Zustände $n \cdot m$. Die Ungewißheit des Verbunduniversums ist dann

$$H(U_1 \,\&\, U_2) = -k \sum_{ij} p_{ij} \log(p_{ij})$$

$$= -k \sum_{ij} p_i \cdot p_j \log(p_i \cdot p_j)$$

$$= -k \sum_{ij} p_i \cdot p_j \log(p_i) - k \sum_{ij} p_i \cdot p_j \log(p_j)$$

$$= -k \sum_{j} p_j \sum_{i} p_i \log(p_i) - k \sum_{i} p_i \sum_{j} p_j \log(p_j).$$

Da $\sum_{1}^{n} p_i = \sum_{1}^{m} p_j = 1$, wird der obige Ausdruck zu

$$H(U_1 \,\&\, U_2) = -k \sum_{1}^{n} p_i \log(p_i) - k \sum_{1}^{m} p_j \log(p_j) = H(U_1) + H(U_2). \quad \text{Q.E.D.}$$

In unserem Unsicherheitsmaß H ist nun nur noch die nach wie vor undefinierte Konstante k in angemessener Weise festzulegen. Für die richtige Wahl dieser Konstante können wir nun wiederum unsere Phantasie spielen lassen. Der heute allgemein anerkannte Vorschlag besteht darin, diese Konstante so zu wählen, daß eine „einfache Einheit der Unsicherheit", gewöhnlich „ein Bit" genannt, mit einem Universum verbunden wird, das aus einer ehrlichen Münze besteht. Die banale Interpretation dieser Wahl besteht darin, daß ein Universum mit genau zwei gleichwahrscheinlichen Zuständen tatsächlich eine gute Norm darstellt, die gewisse elementare Merkmale von Unsicherheit aufweist. Eine etwas kompliziertere Begründung dieser Wahl läßt sich mit Problemen der optimalen Kodierung verknüpfen (Quastler 1958). Ich sehe allerdings im gegenwärtigen Zusammenhang keinen Grund, diese Frage im Detail zu behandeln. Wir wollen daher den zuvor gemachten Vorschlag beibehalten und H den Wert einer Maßeinheit der Unsicherheit geben, die mit einem Universum verknüpft wird, das aus einer ehrlichen Münze besteht:

$$H_{\text{Münze}} = 1 = -k \left[\frac{1}{2} \log \frac{1}{2} + \frac{1}{2} \log \frac{1}{2}\right]$$

$$= k \log 2.$$

Daraus folgt:

$$k = \frac{1}{\log 2},$$

und

$$H = -\frac{1}{\log 2} \sum_{1}^{n} p_i \log p_i,$$

oder, wenn wir 2 zur Basis unserer logarithmischen Skala machen:

$$H = -\sum_{1}^{n} p_i \log_2 p_i. \qquad (6)$$

Mit diesem Ausdruck haben wir das gewünschte Unsicherheitsmaß eines Universums erreicht, das durch n Zustände S_i definiert ist, welche mit der Wahrscheinlichkeit p_i auftreten. Es ist vielleicht angebracht, zur Illustration gewisser Eigenschaften dieser Maßfunktion einige Erläuterungen zu geben. Als erstes möchte ich herausstellen, daß das Unsicherheitsmaß H in einem Universum mit einer festgelegten Anzahl von Zuständen m dann sein Maximum erreicht, wenn alle Zustände mit der gleichen Wahrscheinlichkeit auftreten. Jede Entfernung von dieser einheitlichen Wahrscheinlichkeitsverteilung verringert unmittelbar den Betrag von H, mit anderen Worten, verringert die Unsicherheit des Universums. Lassen Sie mich dies anhand eines Würfels veranschaulichen, der zwar „ehrlich" geboren wurde, aber im Zuge seiner Interaktionen mit schlechter Gesellschaft „korrumpiert" wird. Mein Opfer ist ein Würfel mit einem harten Mantel, der mit einer leimartigen Masse von hoher Viskosität gefüllt ist und im Mittelpunkt eine schwere Stahlkugel enthält. Da diese Anordnung vollkommen symmetrisch ist, nimmt der Würfel bei jedem Wurf mit gleicher Wahrscheinlichkeit einen seiner sechs möglichen Zustände ein. Nun möchte ich aber dem Burschen beibringen, eine Präferenz für die Fläche mit einem Auge zu entwickeln. Zu diesem Zwecke bringe ich unter dem Tisch einen Elektromagneten an und gebe dem Würfel immer dann einen kurzen magnetischen Schock, wenn er die Fläche mit einem Auge zeigt. Dadurch wird die Stahlkugel geringfügig zur unteren Seite hin bewegt, und die Schwerkraft wird das ihre dazu tun, daß er das nächste Mal auf die gleiche Seite fällt. Tabelle 2 zeigt die Wahrscheinlichkeitsverteilung für die verschiedenen Zustände, wie sie sich in den aufeinanderfolgenden Intervallen t_0, t_1, t_2, t_3 und t_4 zwischen den einzelnen magnetischen Schocks darstellt. Auf dem Rand rechts sind die entsprechenden Werte von H eingetragen, d.h. der Unsicherheit des Universums, wie sie der Wahrscheinlichkeitsverteilung der aufeinanderfolgenden Zustände entsprechen; die Werte für H sind in „Bits" angegeben und entsprechend der Gleichung (6) berechnet.

Tabelle 2

Zeit	p_1	p_2	p_3	p_4	p_5	p_6	H Bits
t_0	1/6	1/6	1/6	1/6	1/6	1/6	$2{,}582 = \log_2 6$
t_1	1/4	1/6	1/6	1/6	1/6	1/12	2,519
t_2	1/3	1/6	1/6	1/6	1/6	0	2,249
t_3	2/3	1/12	1/12	1/12	1/12	0	1,582
t_4	1	0	0	0	0	0	0,000

Ein weiteres Merkmal des Unsicherheitsmaßes H besteht darin, daß die Unsicherheit durch den Übergang von einem Universum mit n gleichwahrscheinlichen Zuständen zu einem anderen Universum mit doppelt so vielen gleichwahrscheinlichen Zuständen, 2n, um genau ein Bit zunimmt:

$$H_1 = \log_2 n,$$
$$H_2 = \log_2 2n = \log_2 2 + \log_2 n = 1 + H_1.$$

Ein Universum mit 1 Million Zuständen hat daher eine Unsicherheit von etwa 20 Bits. Fügt man eine weitere Million Zustände hinzu, dann hat das neue Universum etwa 21 Bits Unsicherheit.

Ich habe unsere Maßfunktion H bis jetzt immer als ein Maß der Unsicherheit bezeichnet. Es gibt jedoch mehrere Begriffe, die die gleiche Quantität H bezeichnen, wie sie durch die Gleichung (6) definiert wird. Diese Begriffe sind „Entropie", „(Aus-)Wahl" und „Informationsbetrag".

Die Größe H als Entropie eines Universums oder eines Systems zu bezeichnen, ist durch die Tatsache gerechtfertigt, daß diese thermodynamische Variable, wenn sie durch die Wahrscheinlichkeitsverteilung der ein thermodynamisches System bildenden Moleküle dargestellt wird, durch eine Gleichung definiert wird, die mit unserer Gleichung (6) für H fast identisch ist. Die Bedeutung, die dem Begriff der Entropie gewöhnlich zugemessen wird, ergibt sich aus einer der Konsequenzen des Zweiten Hauptsatzes der Thermodynamik, die besagt, daß die Entropie in einem geschlossenen thermodynamischen System entweder konstant bleibt (thermisches Gleichgewicht) oder steigt, niemals aber abnehmen kann. Dies ist ein Ausdruck der Tatsache, daß die Wahrscheinlichkeitsverteilung der Zustände in „natürlichen Systemen" zur Uniformität hin tendiert, wie durch einen Eimer heißen Wassers in einem kalten Zimmer veranschaulicht werden kann. Nach einer gewissen Zeit hat sich die Wärmeenergie des Eimers mehr oder minder einheitlich über das Zimmer verteilt (Gleichgewicht; alle p_i sind gleich; H ist an seinem Maximum). Ein Thermodynamiker, der nichts von meiner Magnetvorrichtung weiß und beobachtet, wie mein Würfel den Zweiten Hauptsatz der Thermodynamik verletzt, indem er langsam aber sicher von hohen Werten der Entropie zu immer niedrigeren übergeht, wird den Schluß ziehen, daß ein Maxwellscher Dämon am Werk ist, der in selektiver Weise die innere Organisation des Systems verändert. Und wie recht er hat! *Ich* bin natürlich der Dämon, der in selektiver Weise immer dann den Magneten einschaltet, wenn der Würfel die Fläche mit einem Auge zeigt!

Manchmal wird H, wie durch Gleichung (6) definiert, als die Menge der Entscheidungen in einem Universum bezeichnet, das aus n Elementen besteht, und in dem Elemente so ausgewählt werden können, daß einem Element S_i eine Wahrscheinlichkeit p_i zugeordnet wird. Alle die intuitiven Überlegungen, die uns bei der Definition einer Maßfunktion der Unsicherheit geholfen haben – im besonderen das Additionstheorem – lassen sich ebenso gut auf eine Maßfunktion für Entscheidungen anwenden. Die entsprechend abgeleitete Funktion, die ein Maß der Entscheidungen ausdrückt, ist daher identisch mit der Funktion, die ein Maß der Unsicherheit ausdrückt, – sogar hinsichtlich der Maßeinheiten, wenn eine Einheit der Entscheidung mit der gleich starken Versuchung durch zwei Wahlmöglichkeiten gleichgesetzt wird, wie dies Buridans Esel so schön veranschaulicht.

Schließlich wird H auch als der „Informationsbetrag" einer Situation verstanden, in der der gegebene Zustand eines Universums mit der Unsicherheit H von einem Beobachter an einen Empfänger übermittelt wird. Bevor der Empfänger über das Wissen um den faktischen Zustand des Universums verfügt, ist seine Unsicherheit hinsichtlich dieses Universums H. Es stellt sich dann die Frage, welchen Wert er der Information über den Zustand des Universums zumißt, die ihm vom Beobachter übermittelt wird. Auch hier sind all die intuitiven Überlegungen anwendbar, die wir zur Definition eines Unsicherheitsmaßes herangezogen haben, und die sich ergebende Funktion, die ein Maß der Information ausdrückt, ist

identisch mit H, wie durch Gleichung (6) definiert. Da die fraglichen „Zustände" in einer Kommunikationssituation gewöhnlich aus Symbolen bestehen, wird H üblicherweise in Bits pro Symbol ausgedrückt. Wenn der Beobachter Symbole in einer konstanten Abfolge übermittelt, kann H auch durch Bits pro Sekunde ausgedrückt werden.

Wenn wir jedes Wort unseres Wortschatzes – der vermutlich etwa 8000 Wörter umfaßt – mit gleicher Wahrscheinlichkeit benutzten, hätte unser sprachliches Universum eine Unsicherheit H von 13 Bits ($H = \log_2 8192 = 13$). Da wir jedoch die verschiedenen Wörter mit unterschiedlicher Häufigkeit benutzen, ist die Unsicherheit unseres sprachlichen Universums etwas kleiner; sie wird mit etwa 11 Bits angegeben (Shannon 1951). Immer dann also, wenn ich ein Wort ausspreche, übermittle ich im Durchschnitt etwa 11 Bits an Information an Sie. Da ich etwa drei Wörter pro Sekunde äußere, liegt meine Informationserzeugungsrate bei etwa 33 Bits pro Sekunde. (Hoffentlich erzählt mir nun niemand, daß das, was ich erzeuge, nicht Information, sondern bloßes Rauschen ist.)

Ich hoffe, daß ich Ihnen in dieser kurzen Skizze habe zeigen können, daß ein und derselbe Ausdruck, nämlich $-\Sigma\, p_i \log_2 p_i$, ein Maß für ganz verschiedene Größen in ganz verschiedenen Zusammenhängen bezeichnet. Dies spiegelt sich auch in den unterschiedlichen Namen dieses Ausdrucks, wie etwa Unsicherheit, Entropie, Entscheidung, und Information. Diese Sachlage läßt sich auch in der Mechanik aufzeigen, wo das Produkt aus Kraft und Länge(neinheit) in einem Zusammenhang „Arbeit", in einem anderen aber „Drehmoment" bedeutet.

Der nächste Schritt in meiner Darlegung besteht nun darin, diese quantitativen Begriffe zur Konstruktion einer weiteren Maßfunktion zu benutzen, und zwar für ein Maß der „Ordnung". Erneut werde ich mich von intuitiven Überlegungen leiten lassen, um aus der großen Masse möglicher Maßfunktionen diejene auszuwählen, die bestimmte erwünschte Kriterien erfüllt.

Als erstes möchte ich festhalten, daß wir immer dann, wenn wir von „Ordnung" sprechen, diesen Begriff in einem relativen Sinn, d.h. mit Bezug auf den Zustand der Ordnung eines bestimmten Universums gebrauchen. Wir sagen etwa, daß ein *Zimmer* sich in verschiedenen Zuständen der Ordnung oder Unordnung befindet, daß ein Schreibtisch ein Chaos ist usw. Wenn wir also feststellen wollen, daß ein gegebenes „Universum" in vollständiger Unordnung ist, sollte die Funktion, die unser Maß der Ordnung darstellt, verschwinden. Umgekehrt kann daher vollkommene Ordnung durch eine Eins ausgedrückt werden. Verschiedene Ordnungszustände eines gegebenen Universums können folglich durch eine Zahl zwischen 0 und 1 dargestellt werden. Weitere Hinweise auf die allgemeine Form der Funktion, die ein Ordnungsmaß ausdrückt, lassen sich aus der trivialen Tatsache gewinnen, daß unsere Unsicherheit H hinsichtlich eines vollkommen ungeordneten Universums ihr Maximum erreicht ($H = H_{max}$), während die Ordnung in einem deterministischen Universum vollkommen ist ($H = 0$). Dies bedeutet, daß das Ordnungsmaß, welches ein Beobachter mit einem bestimmten Universum verbindet, genau in der Differenz zwischen der für ihn bestehenden und der maximalen Unsicherheit dieses Universums im Verhältnis zum Unsicherheitsmaximum besteht. Wir können daher versuchsweise Ω als die Maßfunktion der Ordnung folgendermaßen definieren:

$$\Omega = \frac{H_{max} - H}{H_{max}} = 1 - \frac{H}{H_{max}} . \qquad (7)$$

Diese Funktion erfüllt klarerweise die beiden oben erörterten Bedingungen, denn bei $H = H_{max}$ wird das Ordnungsmaß verschwinden ($\Omega = 0$), während bei vollkommener Ordnung die Unsicherheit verschwindet und das Ordnungsmaß sich 1 nähert ($\Omega = 1$).

Ich schlage nun vor, diesen Ausdruck für den Additionsfall zu testen. Nehmen wir zwei Universen, U_1 und U_2, für die wir Maße gegebener bzw. maximaler Unsicherheit, H_1, H_{m1} und H_{m2} bestimmen. Wir nehmen weiter an, daß beide Universen im gleichen Ordnungszustand sind:

$$\Omega_1 = 1 - \frac{H_1}{H_{m1}} = \Omega_2 = 1 - \frac{H_2}{H_{m2}}$$

oder

$$\frac{H_1}{H_{m1}} = \frac{H_2}{H_{m2}}.$$

Diese Bedingung ist erfüllt, wenn

$$H_2 = kH_1,$$

und

$$H_{m2} = kH_{m1}.$$

Nun schlage ich vor, die Unterscheidung zwischen den beiden Universen aufzugeben und beide Universen als Teile eines umfassenderen Universums zu behandeln. Worin besteht nun das Maß der Ordnung des umfassenden Universums? Mit Hilfe der Gleichung (7), die Ω definiert, der Gleichung (3), dem Additionstheorem für H, und schließlich mit Hilfe der oben formulierten Identitäten ergibt sich:

$$\Omega = 1 - \frac{H_1 + H_2}{H_{m1} + H_{m2}} = 1 - \frac{H_1(1+k)}{H_{m1}(1+k)} = 1 - \frac{H_1}{H_{m1}} = \Omega_1 = \Omega_2.$$

Mit anderen Worten, die Kombination von zwei gleichermaßen geordneten Universen verändert das Ordnungsmaß nicht, genau wie es intuitiv sein sollte.

Dieses Ordnungsmaß wird sich als hilfreich erweisen, wenn die Veränderungen der inneren Organisation unseres (Er-)Rechners für induktive Schlüsse im Prozeß seiner Interaktion mit der Umwelt quantitativ erfaßt werden sollen. Lassen Sie mich Ihnen nun gleich einige Beispiele von Systemen geben, deren Ordnung auf Kosten externen oder internen Energieverbrauchs zunimmt. Ein Beispiel habe ich bereits gegeben: den magnetischen Würfel. Tabelle 3 verzeichnet die Werte von Ω für die verschiedenen Zeitintervalle und Unsicherheiten aus Tabelle 2.

Tabelle 3

Zeit	t_0	t_1	t_2	t_3	t_4
H	2,582	2,519	2,249	1,582	0,000
Ω	0,000	0,025	0,129	0,387	1,000

Ein weiteres Beispiel wäre das Kristallwachstum in einer übersättigten Lösung. Erneut nimmt die Organisation des Systems zu, indem die diffundierten Moleküle sich an das Kristallgitter binden; in diesem Prozeß wird jedoch ihre potentielle Energie verringert.

PRIBRAM:

Hier ergibt sich ein Problem, denn der Biologe ist weniger an dieser Art von Ordnung interessiert, als an jener, die mit sequentiellen Abhängigkeiten zusammenhängt. Aus diesem Grunde wäre ein hierarchisches Maß irgendeiner Art geeigneter, es sei denn, Sie können zeigen, daß Sie ein solches aus Ihrer Gleichung ableiten können.

von FOERSTER:

Mein Ordnungsmaß ist so allgemein, daß es, so meine ich, keinerlei Schwierigkeiten geben sollte, mit dem Typ von Ordnung fertig zu werden, der sich aus sequentieller Abhängigkeit ergibt. Erlauben Sie mir darzulegen, wie dies meiner Ansicht nach geschehen kann. Sequentielle Abhängigkeiten drücken sich in der Form von Übergangswahrscheinlichkeit p_{ij} aus, d.h. in den n^2 Wahrscheinlichkeiten dafür, daß ein System, welches sich im Zustand S_i befindet, zum Zustand S_j übergeht. Sind die Zustände eines Systems unabhängig von vorausgegangenen Zuständen, wie dies bei der Münze und dem Würfel der Fall ist, sind alle p_{ij}'s natürlich nichts weiter als p_j's. Ein System jedoch, das lernt, wird starke sequentielle Abhängigkeiten entwickeln, wie Sie eben angedeutet haben, und wird folglich die p_{ij}'s von den p_j's abrücken. Erneut kann ein Maß der Unsicherheit H für diesen Sachverhalt einfach dadurch definiert werden, daß man mit dem Mittelwert der verschiedenen Unsicherheiten H_i arbeitet, die für alle Zustände berechnet werden können, die unmittelbar dem Zustand S_i folgen. Da

$$H_i = -\sum_{j=1}^{n} p_{ij} \log_2 p_{ij},$$

lautet unsere Regel für die Ableitung eines Mittelwertes:

$$H = \overline{H}_i - \sum_{1}^{n} p_i H_i.$$

Wenn die Anzahl der Zustände des Systems n gleich bleibt, ändert sich auch $H_{max} = \log_2 n$ nicht; wir können dagegen beobachten, daß Ω ständig zunimmt, wenn die sequentiellen Beschränkungen wachsen — d.h. wenn H abnimmt —, während das Lebewesen trainiert wird.

Da die Veränderung der inneren Organisation eines Systems, ob im räumlichen oder zeitlichen Sinn, zu immer höheren Ebenen der Organisation ein Kernpunkt meiner Beschreibung der sogenannten „selbstorganisierenden Systeme" ist, jener Systeme also, die Umweltordnung in ihrer eigenen Organisation abbilden, möchte ich nun die Kriterien festlegen, die erfüllt werden müssen, wenn wir unsere Systeme zu derartigen selbstorganisierenden Systemen machen wollen. Klarerweise sollte Ω bei solchen Systemen im Verlauf der Zeit zunehmen:

$$\frac{d\Omega}{dt} > 0.$$

Da unser Maß der Ordnung eine Funktion von H und H_{max} ist, die beide Veränderungen unterworfen sein können oder auch nicht, erreichen wir das gewünschte Kriterium durch die Differenzierung der Gleichung (7) mit Bezug auf die Zeit[1]):

$$\frac{d\Omega}{dt} = -\frac{H_m \frac{dH}{dt} - H \frac{dH_m}{dt}}{H_m^2} > 0.$$

Dieser Ausdruck läßt sich in eine etwas konkretere Form transformieren. Wir bedenken zuerst, daß für alle uns interessierenden Systeme $H_m > 0$ ist, da nur bei Systemen, die genau einen Zustand einnehmen können, $H_m = \log_2 1 = 0$. Dann dividieren wir beide Seiten der Ungleichung durch das Produkt $H \cdot H_m$ und erhalten so die wichtige Beziehung:

$$\frac{1}{H_m} \frac{dH_m}{dt} > \frac{1}{H} \frac{dH}{dt}. \tag{8}$$

Damit wird gesagt, daß sich unser System dann und nur dann, wenn die relative Zunahme der maximalen Unsicherheit größer ist als die relative Zunahme der faktischen Unsicherheit, in einem Prozeß befindet, höhere Zustände interner Organisation aufzubauen.

BOWER:

Spielen dabei empirische Überlegungen eine Rolle?

von FOERSTER:

Nein, überhaupt keine. Es handelt sich hier um eine direkte Ableitung auf der Basis einer Definition und mit Hilfe eines Kriteriums.

BOWER:

Ich denke an mehrere Gegenbeispiele, bei denen das Verhalten zunehmend zufallsgesteuert oder in Ihrem Sinne ungeordnet wird, obwohl etwas gelernt wird. Ich bin z.B. sicher, daß ich eine Ratte trainieren könnte, ihr Verhalten von einem Versuch zum anderen so zu variieren, daß es für mich völlig unvorhersagbar wird: Man verstärkt ganz einfach in differentieller Weise die Varianz entsprechend einem bestimmten Kriterium – z.B. dadurch, daß das Niederdrücken eines Hebels durch die Ratte nur dann bekräftigt wird, wenn seine Latenz sich um zumindest zwei Sekunden vom vorausgegangenen Niederdrücken des Hebels unterscheidet. Des weiteren nimmt die Referenzreaktion bei der Auslöschung gelernten Verhaltens in ihrer Stärke ab und tritt mit geringerer Sicherheit auf. Sie würden dies als zunehmende Unordnung beschreiben, doch handelt es sich dabei um einen gesetzmäßigen und einheitlichen Prozeß, besonders dann, wenn konkurrierende Reaktionen aufgezeichnet werden.

von FOERSTER:

Ich würde sagen, daß das Training in einer derartigen Situation die Verhaltensmöglichkeiten für Ratten vermehrt, indem es im Verhaltensuniversum der Ratten neue Zustände erzeugt. Sie beeinflussen daher H_{max} so, daß dH_m/dt größer

[1]) H_{max} wird hier aus typographischen Gründen als H_m geschrieben.

ist als 0. Verschlechtert sich der Zustand dieser Tiere nicht in irgendeiner anderen Hinsicht, und nimmt H folglich nicht zu schnell zu, dann haben Sie diesen Ratten in der Tat etwas beigebracht.

JOHN:

Könnten Sie bitte H_{max} definieren?

von FOERSTER:

H_{max} kann einfach als das Unsicherheitsmaß eines Systems mit gleichwahrscheinlichen Zuständen definiert werden, die von ihren vorausgegangenen Zuständen unabhängig sind. Unter diesen Bedingungen ist, wie wir schon gesehen haben, H_{max} genau $\log_2 n$, wobei n die Zahl der Zustände bezeichnet.

JOHN:

Genau. Mir scheint daher die Ableitung von H_m mit Bezug auf die Zeit stets 0 sein zu müssen.

von FOERSTER:

Das ist ein ausgezeichneter Hinweis. Sie haben damit auf ein faszinierendes Merkmal dieser Gleichung hingewiesen, darauf nämlich, daß sie Wachstum erklärt. Damit haben Sie bereits das nächste Kapitel meiner Geschichte vorweggenommen.

Um unmittelbar klarzumachen, daß d_{Hm}/dt nicht notwendig immer 0 sein muß, will ich H_m durch $\log_2 n$ ersetzen, oder der Einfachheit halber durch $\ln n$, wobei $a = 1/\ln 2$ ein skalarer Faktor ist, der die Logarithmen zur Basis 2 in natürliche Logarithmen transformiert. Daraus ergibt sich:

$$\frac{dH_m}{dt} = a\frac{d\ln n}{dt} = \frac{a}{n}\frac{dn}{dt}. \tag{9}$$

Stellen Sie sich einen Augenblick einen Organismus vor, der durch Zellteilung wächst. In den frühen Phasen seiner Entwicklung wächst die Anzahl seiner Zellen gewöhnlich exponentiell:

$$n = n_0 e^{\lambda t},$$

und aufgrund von

$$H_m = a \ln n = a \ln n_0 + a\lambda t, \tag{10}$$

wird die Rate der Veränderung der maximalen Entropie zu

$$\frac{dH_m}{dt} = a\lambda,$$

und dies ist eine positive Konstante! Das bedeutet, daß die tatsächliche Entropie H des Systems nicht notwendig abnehmen muß, damit das System höhere Zustände der Organisation erreicht. Nach Gleichung (8) genügt es, daß die relative Veränderungsrate von H nur unterhalb jener von H_{max} bleibt. Ein Beobachter, der lediglich die Zunahme der faktischen Unsicherheit H beachtet, mag dem Eindruck verfallen, daß sein System kaputt geht. Wenn wir unsere Städte heute anschauen, können wir leicht diesen Eindruck gewinnen. Wenn wir jedoch andererseits ihr rapides Wachstum betrachten, erweisen sie sich als Zentren zunehmender Organisation. Aber lassen Sie mich zur Population sich selbst teilender und dif-

ferenzierender Zellen zurückkommen. Das organisierte Wachstum von Gewebe unterliegt beträchtlichen Einschränkungen der möglichen Anordnung der Zellen, und die Wahrscheinlichkeitsverteilung ihrer Positionen ist daher weit von jeder Gleichverteilung entfernt. H verändert sich folglich im Verlauf der Wachstumsphase des Organismus tatsächlich sehr langsam, wenn überhaupt. Nichtsdestoweniger läßt sich das folgende konservative Modell für eine versuchsweise Formulierung des Wachstums von H nach der Gleichung (10) geben:

$$H = a \ln n_0 + \mu a \lambda t, \tag{11}$$

wobei $\mu < 1$. Das Maß der Ordnung im wachsenden Organismus wird mit Hilfe der Gleichungen (10) und (11) zu

$$\Omega = 1 - \frac{\ln n_0 + \mu\lambda t}{\ln n_0 + \lambda t}.$$

In den frühen Stadien seiner Entwicklung ($t = 0$) ergibt sich

$$\Omega(0) = 0,$$

dagegen im Reifezustand ($t \to \infty$):

$$\Omega(\infty) = 1 - \mu > 0;$$

d.h. der Organismus ist in der Tat ein „Organismus".

Die Möglichkeit, den Erwerb höherer Zustände der Organisation durch die Eingliederung neuer Zustände in das System zu erklären — *vorausgesetzt*, daß diese Eingliederung in geordneter Weise erfolgt —, ist hier für mich durch die schönen Arbeiten verdeutlicht worden, die Dr. Hydén vorgetragen hat. Nehmen wir einmal den Kern eines Neurons als unser Universum. Die Zunahme an Organisation kann ganz ungeheuer sein, wie sich aus der Gleichung (9) ersehen läßt, die den absoluten Betrag der maximalen Unsicherheit der perzentuellen Zunahme der Zahl geordneter Elemente proportional setzt. Da es in diesem Kern nur 80000 Moleküle gibt, können bereits ein paar tausend Moleküle, die nach Dr. Hydéns kompliziertem Mechanismus modifiziert werden, wesentlich zum Ω des Systems beitragen.

Ich habe bisher nur etwas allgemeine Bemerkungen über die Nützlichkeit von Zahlen bei der Beschreibung von Systemen gemacht, die sich in unterschiedlichen Zuständen der Ordnung befinden, von Zahlen, die einen Betrag an Unsicherheit, an Komplexität, oder an „Perplexität" solcher Systeme ausdrücken (wie es einer meiner Studenten formulierte), ebenso wie schließlich den Betrag der Information, der benötigt wird, um sie im einzelnen zu entwickeln. Ich möchte nun zum Thema unserer Tagung zurückkehren, für das diese Zahlen nützlich werden können.

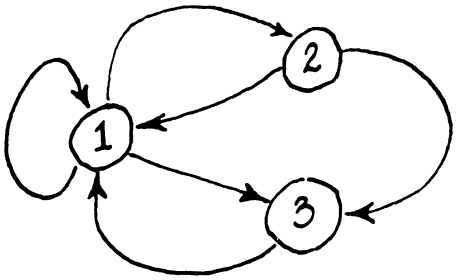
Zur Diskussion steht eine wichtige Eigenschaft des Funktionierens unseres Nervensystems. Wir nennen sie „Gedächtnis". Was die Suche nach Mechanismen angeht, die für diese Eigenschaft verantwortlich gemacht werden können, habe ich mit Nachdruck betont, daß wir dieses System nicht als eine Art von Aufzeichnungsvorrichtung betrachten sollten. Ich habe stattdessen vorgeschlagen, dieses System als eine Art von Rechner anzusehen, dessen innere Organisation sich aufgrund seiner Interaktionen mit einer Umwelt, die eine gewisse Ordnung aufweist,

verändert. Die Veränderungen der inneren Organisation dieses Rechners finden so statt, daß bestimmte Gesetzlichkeiten der Umwelt, die für deren Ordnung verantwortlich sind, in der Struktur dieses Rechners abgebildet werden. Diese Homomorphie „Umwelt-System" ist das „Gedächtnis", es erlaubt dem System, als ein (Er-)Rechner induktiver Schlüsse zu arbeiten. Umweltzustände, die sozusagen „mit Gesetzen der Natur unvereinbar" sind, sind auch mit Outputzuständen des Rechners unvereinbar.

Ich möchte nun die Numerik der Informationstheorie auf einige der bekannten Merkmale des Nervensystems anwenden, um zu zeigen, welche Schlußfolgerungen sich aus diesen Zahlen ableiten lassen. Die erste Zahl, die ich ableiten will, besteht in einem Schätzwert des Informationsbetrages, der zur Bestimmung eines „Gehirns" notwendig ist. Um nun bei derartigen Schätzungen überhaupt voranzukommen, muß man, wie ich bereits früher dargelegt habe, zuerst das „Universum" – in unserem Falle das Gehirn – durch eine endliche Anzahl von Zuständen und die Wahrscheinlichkeiten ihres Auftretens im einzelnen festlegen. Ist dies geschehen, dann kann H (Gehirn) aus der Gleichung (7) berechnet werden. Ich schlage daher vor, „Gehirn" als die Menge einer endlichen Anzahl von Elementen, den Neuronen nämlich, aufzufassen, die auf bestimmte Art und Weise miteinander verbunden sind und ein riesiges Netzwerk bilden. Ich schlage vor, diesen Verbindungen „Richtung" zu geben, indem ich die Verbindungslinien mit imaginären Pfeilen ausstatte, die die Unidirektionalität der Ausbreitung von Nervenimpulsen entlang der Axone ausdrücken, die vom Zellkörper wegführen. Das zu bearbeitende Universum besteht folglich aus allen möglichen Netzwerken, die sich durch die Verbindung der Elemente bilden lassen, und jeder einzelne Zustand dieses Universums ist ein bestimmtes Netzwerk. Ich muß nun die Anzahl der Zustände dieses Universums abschätzen, mit anderen Worten, die Anzahl der verschiedenen Netzwerke, die sich dadurch bilden lassen, daß ich n Elemente in richtungsspezifischer Weise verbinde. Die Frage, was ein Netzwerk von einem anderen unterscheidet, kann von zwei verschiedenen Standpunkten angegangen werden. Der eine ist ein rein struktureller und vernachlässigt die operationalen Modalitäten der Verknüpfungselemente; der andere berücksichtigt eben diese Modalitäten. Ich schlage vor, zuerst die rein strukturellen Merkmale zu betrachten und erst später die möglichen Arbeitsweisen der einzelnen Neuronen in ihren strukturell definierten Netzwerken.

Das Problem, die Anzahl der Netze abzuzählen, die dadurch gebildet werden können, daß n Elemente in richtungsspezifischer Weise verknüpft werden, wird mit Hilfe einer Verbundmatrix leicht gelöst. Diese Matrix besteht aus einem Quadrat von n Zeilen und n Spalten, die mit den Namen der entsprechenden Elemente versehen sind (Bild 1). Wird Element E_i mit Element E_j verknüpft, wird in die i-te Zeile am Schnittpunkt mit der j-ten Spalte „1" eingetragen, sonst „0". Die spezifische Verteilung von „1" und „0" in der Matrix legt das entsprechende Netzwerk eindeutig fest. $\mathcal{N}$, die Anzahl der möglichen Eintragungen von „1" und „0" in die n^2 Stellen der Matrix, ist daher auch die Anzahl der verschiedenen Netze, die aufgebaut werden können, indem man n Elemente in richtungsspezifischer Weise verknüpft. Da für jede Stelle zwei Möglichkeiten bestehen, ist diese Zahl

$$\mathcal{N} = 2^{n^2}.$$



NET

	1	2	3
1	1	1	1
2	1	0	1
3	1	0	0

MATRIX

Bild 1
Darstellungen von Netzwerken:
(a) Graph;
(b) Matrix. Beide Darstellungen sind äquivalent.

Da ich überhaupt keine Ahnung habe, ob ein Netz wahrscheinlicher ist als ein anderes, ist mein Universum von gleichwahrscheinlichen Zuständen bevölkert, und seine Unsicherheit ist die folgende:

$$H = \log_2 2^{n^2} = n^2 \text{ Bits/Netz.}$$

Mit anderen Worten, n^2 Bits an Information sind notwendig, um ein bestimmtes Netzwerk festzulegen, wie auch direkt aus der Verbundmatrix hätte abgelesen werden können, wo es n^2 binärer Entscheidungen bedurfte, „1" oder „0" in die n^2 Stellen einzutragen, um ein einzelnes Netzwerk festzulegen.

Die Schätzungen der Anzahl der Neuronen in einem menschlichen Gehirn kreisen um etwa 10 Milliarden. Daraus folgt:

$$H = (10^{10})^2 = 10^{20} \text{ Bits/Gehirn.}$$

Wir wollen nun prüfen, ob die Information, die benötigt wird, um die Verknüpfungsstruktur des Nervensystems festzulegen – von der Bestimmung der operationalen Modalitäten seiner Elemente abgesehen – genetisch determiniert werden kann oder nicht. Glücklicherweise gibt es gute Schätzungen des Informationsgehaltes des genetischen Programms. Die sorgfältigste ist meiner Meinung nach immer noch die von Dancoff und Quastler (1953). Diese Autoren gelangen aufgrund verschiedener Überlegungen zu einer Ober- und einer Untergrenze für den Betrag der Unsicherheit H_G in einer einzelnen Zygote:

$$10^5 < H_G < 10^{12}.$$

Mit anderen Worten, das Programm, das vermutlich die Struktur definiert, liegt um einen Faktor von etwa 10^{10} von der geforderten Größenordnung entfernt. Dies deutet ganz klar darauf hin, daß der genetische Kode, der weitaus mehr festlegt als nur das Nervensystem, nicht in der Lage ist, Netzwerke der unbeschränkten Allgemeinheit zu programmieren, wie wir sie vorhin betrachtet haben. Ein Weg aus diesem Dilemma besteht darin anzunehmen, daß faktisch nur ein außerordentlich kleiner Anteil der Struktur des Nervensystems genetisch festgelegt wird, während der überwältigende Teil dem Zufall überlassen bleibt. Auch wenn man die Vorstellung, daß für Zufallsverknüpfungen ein gewisser Raum sein muß, nicht völlig verwerfen kann, scheint es dennoch wenig einleuchtend, daß nur ein Hundertstel von 1 % oder weniger aller Neuronen festgelegte Verknüpfungsstrukturen aufweisen. Diese Annahme würde zum Beispiel jede Neuroanatomie unmöglich machen, denn die Verschiedenheiten der Gehirne würden ihre Ähnlichkeiten bei weitem übertreffen.

Ein anderer Weg aus diesem Dilemma besteht in der Annahme, daß der genetische Kode in der Tat in der Lage ist, eine große Vielfalt von Netzwerken zu programmieren, wobei jedoch jedes Netzwerk nur eine kleine Anzahl von Neuronen umfaßt und in paralleler Weise immer wieder wiederholt wird. Die Wiederholungen einer bestimmten Struktur in paralleler Weise erfordert sehr wenig Information, denn der einzige notwendige Befehl lautet „Wiederhole diese Operation bis Stop!" Die verschiedenen Arten der Netzwerke können dann in Form einer Kaskade gestapelt werden (Bild 2). Ganz grob gesprochen, liegt die Attraktivität dieses Bildes darin, daß es gewisse Ähnlichkeiten mit den Lamellenstrukturen aufweist, wie sie die Verteilung der Neuronen in den äußeren Lappen des Gehirns

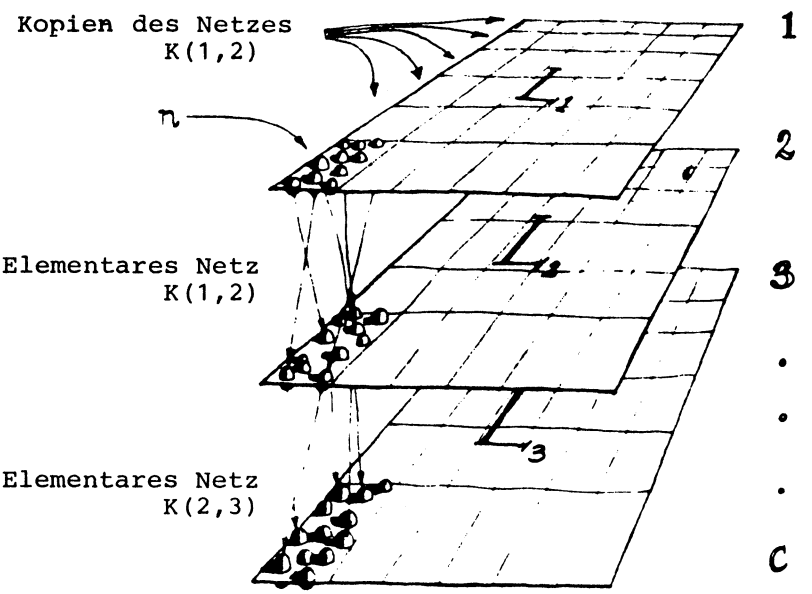


Bild 2 Kaskade von C ungleichen Netzwerken, in paralleler Weise aufgebaut aus P gleichen Netzen.

kennzeichnen. Wir wollen einmal prüfen, welche Zahlen wir bekommen, wenn wir annehmen, daß das ganze System von in Kaskaden angeordneten parallelen Netzwerken durch das genetische Programm festgelegt wird.

Ich schlage vor, ein kleines elementares Netz zu betrachten, das lediglich 2_n Neuronen enthält, deren eine Hälfte in einer Schicht, sagen wir L_1, und deren andere Hälfte in einer angrenzenden Schicht L_2 lokalisiert ist (Bild 2). Die Axone der Neuronen von L_1 verbinden sich mit jenen in L_2; wir nehmen jedoch an, daß es in diesem einfachen Modell keine zurückführenden Pfade gibt. Da die Gesamtanzahl der Neuronen, die in jeder Schicht lokalisiert sind, als sehr groß angenommen wird, sagen wir N, wird das vollständige Verknüpfungsmodell für die beiden Schichten durch eine Verschiebung des elementaren Netzwerks parallel zu sich selbst in beide Richtungen entlang der Oberfläche der Schichten erreicht. Die Anzahl paralleler Netzwerke ist daher

$$P = \frac{N}{n}.$$

Wiederum läßt sich eine Verbundmatrix mit n Zeilen und n Spalten für das elementare Netzwerk zeichnen, die den Neuronen in den Schichten L_1 bzw. L_2 entspricht und in der an jedem Schnittpunkt einer Zeile und einer Spalte das Vorhandensein oder Fehlen einer Verbindung zwischen einem Neuron in der Schicht L_1 und einem Neuron in der Schicht L_2 durch „0" oder „1" angezeigt wird. Die Anzahl der Netze ist folglich wiederum 2^{n^2}, und die Unsicherheit dieses elementaren Netzwerks ist daher

$$H_n = n^2.$$

Auch wenn P derartige Netze parallel zwischen den Schichten L_1 und L_2 arbeiten, ist die Unsicherheit für das gesamte Netzwerk, das diese zwei Schichten verknüpft, lediglich n^2, da es keinerlei Freiheit auch nur für eine einfache Verknüpfung in irgendeinem der P Netzwerke gibt, sich ohne entsprechende Veränderungen in all den übrigen Netzen selbst zu verändern; ihre Verknüpfungsstruktur ist nämlich durch die Verbundmatrix determiniert, die als eine Art genetischer Schablone funktioniert, mit der alle P Netze gebildet werden.

Ich schlage nun vor anzunehmen, daß eine weitere Verbundmatrix die Verbindungen zwischen dem nächsten Schichtenpaar (L_2, L_3) bestimmt usw. in der ganzen Kaskade von C Schichten. Die Unsicherheit des Systems als ganzen ist daher

$$H_S = nC^2,$$

und sie soll wieder durch den genetischen Kode bestimmt sein. Daraus folgt:

$$H_S = H_G.$$

Wir müssen nun aber in diesem System insgesamt $\overline{N}$ Neuronen unterbringen, die über C Schichten verteilt sind, deren jede nP Neuronen enthält:

$$\overline{N} = nPC.$$

Wenn wir n, die Anzahl der Zellen in einem elementaren Netzwerk, aus den zwei oben gegebenen Gleichungen eliminieren, erhalten wir eine Beziehung zwi-

schen der Anzahl der Kaskaden und der Anzahl der parallelen Kanäle in jeder Kaskade:

$$H_G = \frac{\overline{N}^2}{CP^2},$$

oder

$$p = \frac{\overline{N}}{H_G} \cdot \frac{1}{C}.$$

Tabelle 4 bietet für drei denkbare Werte genetischer Information H_G eine Menge von fünf Werten je Dreiergruppe C, P, n, die die obige Gleichung erfüllen.

Mit Bezug auf die verschiedenen Wahlmöglichkeiten, die durch Tabelle 4 ausgedrückt werden, scheint mir für eine angenommene genetische Information von 10^8 Bits pro Zygote ein System, das im Durchschnitt aus 1000 Schichten ($C = 10^3$) besteht, deren jede 30000 parallele elementare Netzwerke ($P = 3 \cdot 10^4$) von jeweils 300 Neuronen enthält, sehr vergröbert eine Strukturskizze der kortikalen Organisation darzustellen, die vielleicht nicht sofort als quantitativ völlig abwegig verworfen werden muß.

Tabelle 4

H_G Bits								
10^6			10^8			10^{10}		
C	P	n	C	P	n	C	P	n
10^2	$1 \cdot 10^6$	100	10^2	$1 \cdot 10^5$	1000	10^2	$1 \cdot 10^4$	10^4
10^3	$3 \cdot 10^5$	30	10^3	$3 \cdot 10^4$	300	10^3	$3 \cdot 10^3$	3000
10^4	$1 \cdot 10^5$	10	10^4	$1 \cdot 10^4$	100	10^4	$1 \cdot 10^3$	1000
10^5	$3 \cdot 10^4$	3	10^5	$3 \cdot 10^3$	30	10^5	$3 \cdot 10^2$	300
10^6	$1 \cdot 10^4$	1	10^6	$1 \cdot 10^3$	10	10^6	$1 \cdot 10^2$	100

Dieses Bild ist natürlich äußerst grob, sein Wert – wenn es überhaupt einen Wert hat – liegt aber vor allem darin, daß es aus der riesigen Menge möglicher Merkmale gewisse hervorhebt, die genauere Betrachtung verdienen.

Ich möchte nun auf einige Konsequenzen zu sprechen kommen, die mir für unser Thema von Bedeutung zu sein scheinen. Erstens einmal erlaubt uns die Möglichkeit paralleler Kanäle, mit relativ kleinen Netzen zu arbeiten, für die einmal eine adäquate Theorie konstruiert werden kann. Ich werde gleich über den gegenwärtigen Wissensstand in der Theorie kleiner Rechennetze berichten. Zweitens errechnet ein Netzwerk, das aus der periodischen Wiederholung desselben elementaren Netzwerks besteht, für alle seine Stimuli die gleichen Funktionen, unabhängig von der linearen Übertragung der Stimulusverteilung. Daraus folgt, daß Parallelismus Invarianz der Übertragung bedingt.

Ich habe eben das elementare Netzwerk als „Rechennetz" bezeichnet. Ich schulde Ihnen nun eine Erklärung dafür, daß dies ein angemessener Begriff ist. Gleichzeitig kann ich zeigen, was solche Netzwerke berechnen, worin ihre Rechen-

möglichkeiten als Bestandteile großer paralleler Netze bestehen, und schließlich, wie sie ihre operationalen Modalitäten aufgrund der Ergebnisse früherer Berechnungen modifizieren können.

Wir haben an der University of Illinois seit 1958 das Rechenpotential periodischer Netzwerke untersucht. Diese Arbeiten sind durch verschiedene Forschungsergebnisse von Lettvin u.a. (1959), Maturana/Frenk (1963), Mountcastle (1963), Hubel/Wiesel (1959; 1962) und anderen gefördert worden, was die Errechnung von Abstraktionen in kleinen, parallel arrangierten Netzen betrifft. Die Grundidee (Babcock u.a. 1960; Inselberg/v. Foerster 1962) besteht dabei darin, das Verknüpfungsschema zwischen zwei Schichten, die gleichmäßig verteilte Rechenelemente enthalten, in geometrische Begriffe zu fassen (Bild 3).

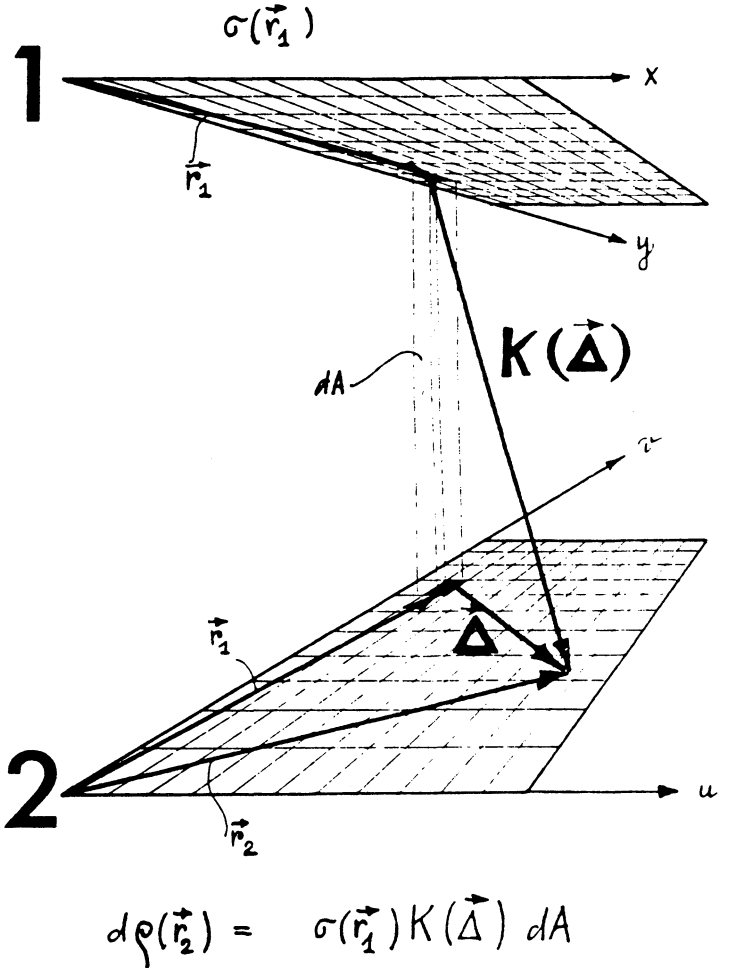


Bild 3 Geometrische Beziehungen in einem Aktionsnetzwerk.

Nehmen wir an, daß sich Nervenfasern von einem kleinen Gebiet dA, das in der Schicht L_1 in $\mathbf{r}_1$ lokalisiert ist, in allen Richtungen nach unten erstrecken, um mit Elementen Synapsen zu bilden, die sich in der Schicht L_2 befinden. Betrachten wir einmal das Bündel, das mit Elementen in $\mathbf{r}_2$ der Schicht L_2 Synapsen bildet. Über dieses Bündel läuft ein gewisser Teil λ der Aktivität $\sigma(\mathbf{r}_1)\,dA$, die in der Nachbarschaft von $\mathbf{r}_1$ herrscht, und ruft eine infinitesimale Reaktion $d\rho(\mathbf{r}_2)$ der Elemente hervor, die in der Nachbarschaft von $\mathbf{r}_2$ in der Schicht L_2 lokalisiert sind. Diese Reaktion sei der Stimulusaktivität in $\mathbf{r}_1$ proportional:

$$d\rho(\mathbf{r}_2) = \lambda K(\mathbf{r}_1, \mathbf{r}_2)\,\sigma(\mathbf{r}_1)\,dA,$$

wobei $K(\mathbf{r}_1, \mathbf{r}_2)$ die Proportionalitäts-„konstante" ist, die von Punkt zu Punkt in der Stimulusschicht L_1 ebenso wie in der Reaktionsschicht L_2 verschiedene Werte annehmen kann. Dies sei dadurch angezeigt, daß K eine Funktion der Orte $\mathbf{r}_1$ und $\mathbf{r}_2$ ist. Da K in eindeutiger Weise mit dem Fasernbündel verknüpft wird, das die Elemente in $\mathbf{r}_1$ mit jenen in $\mathbf{r}_2$ verknüpft, ist K der Parameter, der die Aktivität des elementaren Netzwerks repräsentiert, das die beiden Schichten L_1 und L_2 miteinander verbindet. Von nun an werde ich K als die „Aktionsfunktion" des elementaren Netzwerks bezeichnen. Dieses Netzwerk soll sich außerdem mit der Periodizität p in beiden Richtungen wiederholen, so daß gilt

$$K(x + ip, y + jp; u + ip, v + jp) = K(x, y; u, v)\ i, j = 0; \pm 1; \pm 2; \pm 3; \ldots$$

und folglich K sich nur als Funktion der Entfernung

$$\Delta = \mathbf{r}_2 - \mathbf{r}_1$$

zwischen den zwei betrachteten Punkten darstellt:

$$K(\mathbf{r}_1, \mathbf{r}_2) = (\mathbf{r}_2 - \mathbf{r}_1) = K(\Delta)$$

Die Reaktion in $\mathbf{r}_2$, die durch die Einwirkung aus $\mathbf{r}_1$ erzeugt wird, ist natürlich nur ein Bruchteil der in $\mathbf{r}_2$ hervorgerufenen Reaktion. Um in $\mathbf{r}_2$ eine Totalreaktion zu erreichen, müssen wir die elementaren Einwirkungen aus allen Bereichen der Stimulusschicht zusammenzählen:

$$\rho(\mathbf{r}_2) = \int_{L_1} K(\Delta)\,\sigma(\mathbf{r}_1)\,dA. \tag{13}$$

Wenn die Aktionsfunktion $K(\Delta)$ definiert ist – und mein Vorschlag bestand darin, daß sie durch das genetische Programm festgelegt wird –, dann wird für eine gegebene Stimulusverteilung die Reaktionsverteilung durch den obigen Ausdruck bestimmt. Die physiologische Bedeutung der Aktionsfunktion kann am besten dadurch beleuchtet werden, daß diese Funktion in ein Produkt aus zwei Teilen auseinandergelegt wird:

$$K(\Delta) = D(\Delta) \cdot T(\Delta), \tag{14}$$

wobei $D(\Delta)$ eine Veränderung der Dichte der Fasern bezeichnet, die in $\mathbf{r}_1$ auftreten und in $\mathbf{r}_2$ zusammentreffen. $D(\Delta)$ ist folglich ein *struktureller* Parameter. $T(\Delta)$ beschreibt andererseits die lokalen Transferfunktionen für Fasern, die in $\mathbf{r}_1$ entstehen und mit Neuronen in der Nachbarschaft von $\mathbf{r}_2$ Synapsen bilden. $T(\Delta)$ ist daher ein *funktionaler* Parameter.

Ich hoffe nun, daß ich nicht allzulange die Darstellung dessen aufgeschoben habe, was von diesen Netzen berechnet wird. Leider kann ich keine detaillierte Darstellung der unterschiedlichen Rechenergebnisse liefern, die durch die Analyse verschiedener Aktionsfunktionen K erzielt werden können. Vielleicht aber kann ich zumindest sagen, daß die Ergebnisse dieser Berechnungen Invarianten, oder Abstraktionen, der Stimulusverteilung sind. Ich habe Invarianz bereits als den primären Vorteil des parallelen Einsatzes von Netzwerken von der Übertragung abgehoben. Es dürfte außerdem nicht allzu schwierig sein, sich die Art der Abstraktionen vorzustellen, die berechnet werden, wenn die Aktionsfunktion K gewisse Symmetrieeigenschaften besitzt. Betrachten wir z.B. den Fall, daß die drei Grundtypen der Symmetrie für drei Typen von Aktionsfunktionen gelten: die symmetrische, die antisymmetrische und die zirkulärsymmetrische Aktionsfunktion wird jeweils wie folgt definiert:

$$K_s(\Delta) = K_s(-\Delta),$$
$$K_a(\Delta) = -K_a(-\Delta),$$
$$K_c(\Delta) = K_c(|\Delta|).$$

Klarerweise verleiht K_s den Umkehrungen von Stimuli Invarianz, die zu den Achsen $y = 0$ und $x = 0$ symmetrisch sind, d.h. die Ziffer 3 wird zu ϵ oder ein M zu W; K_a andererseits verleiht Umkehrungen von Stimuli Invarianz, die symmetrisch sind zu den Zeilen $y = x$, d.h. $\backsim$ wird zu S, und $>$ zu V. K_c schließlich verleiht Rotationen Invarianz, d.h. fast allen vorhin genannten Umkehrungen einschließlich N zu Z.

Vielleicht haben Sie in einigen der Merkmale dieser Aktionsfunktionen Ähnlichkeiten mit Eigenschaften entdeckt, wie sie Hubel und andere in den Reaktionsmustern des von ihnen so genannten „Rezeptiven Feldes" beobachtet haben. In der Tat besteht eine sehr enge Beziehung zwischen diesen beiden Begriffen, denn die Kenntnis des „Bereichs" der Aktionsfunktion, d.h. der Zellen in der Zielschicht, die von einer einzigen Zelle in der Ursprungsschicht „gesehen" werden, ermöglicht unmittelbar die Bestimmung des Bereichs der „Rezeptorfunktion", d.h. der Zellen in der Ursprungsschicht, die von der einzelnen Zelle in der Zielschicht gesehen werden. Sei $G(\Delta)$ die Rezeptorfunktion, dann ergibt sich

$$G_s = K_s; \quad G_a = -K_a; \quad G_c = K_c.$$

An diesem Punkt meiner Ausführungen ließe sich nun einwenden, daß all dies mit wohlbekannten Filteroperationen des Nervensystems zu tun hat, was aber hat es mit dem Gedächtnis zu tun? Es trifft zu, daß ich bis jetzt in der Tat versucht habe, einen exakten Rahmen zu entwickeln, innerhalb dessen wir diese Filteroperationen erörtern können. Was aber dann, wenn diese sich aus Interaktionen mit der Umwelt ergeben? Wie würden wir die Errechnung einer Invariante interpretieren, wenn sie sich aus vergangener Erfahrung ergibt, oder wenn sie das Ergebnis eines Lernprozesses ist? Ich wage zu sagen, daß wir eine derartige adaptive Berechnung von Abstraktionen präzise als das Funktionieren eines „Gedächtnisses" interpretieren würden, das als Antwort auf die Befragung durch einen Strom von Informationen eine knappe Kategorisierung liefert. Meine Erinnerung von „Karl Pribram" kann nicht darin bestehen, daß ich meine früheren Aufzeichnungen aller Einstellungen, Gesten usw. befrage, die ich von ihm gespeichert habe. Die Wahrscheinlichkeit, daß ich darunter jene finde, denen ich in diesem Moment gegen-

überstehe, ist nahe Null. Höchstwahrscheinlich gibt es sie überhaupt nicht. Was ich meiner Meinung nach eingebaut habe, ist ein Netz von Rechnern, das aus der riesigen Informationseingabe, die durch mein visuelles System gepumpt wird, all das errechnet, was „pribramisch" ist – d.h. alle jene Kategorien, die ihn und nur ihn definieren –, und die sodann trocken einen kurzen Namen für alle diese Kategorien äußern: „Karl Pribram".

Um nun diese These zu stützen, muß ich zeigen, daß es in der Tat möglich ist, die Art von Rechnernetzwerken, die ich erörtert habe, so zu trainieren, daß sie ihre Rechengewohnheiten von der Errechnung des einen Typs von Abstraktionen zur Errechnung anderer Abstraktionen ändern. In der Sprache der Theorie dieser Rechennetzwerke muß ich zeigen, daß die Aktionsfunktion K, die die errechnete Invariante eindeutig festlegt, nicht notwendig eine unveränderbare Entität ist, sondern unter dem Einfluß verschiedener Wirkgrößen variieren kann. Mit anderen Worten, die temporale Variation von K löst sich nicht notwendig auf:

$$\delta K \neq 0.$$

Bevor ich die Einzelheiten meiner Demonstration darlege, muß ich bekennen, daß wir meines Wissens heute noch weit entfernt sind von einer befriedigenden Theorie adaptiver abstrahierender Netzwerke. Die mathematischen Probleme, auf die man dabei sofort stößt, sind grundlegender Art, und es gibt zur Zeit noch wenig Literatur, die hier weiterhelfen würde. Ich werde also folglich nicht in der Lage sein, spektakuläre Ergebnisse vorzulegen. Andererseits hoffe ich, daß die beiden folgenden einfachen Beispiele explizit genug sind, um zumindest die wesentlichen Konturen des Ansatzes zu verdeutlichen.

In meiner kurzen Skizze einiger Merkmale paralleler Netzwerke habe ich betont, daß das genetische Programm unter bestimmten Bedingungen ausreichen kann, alle Netzwerke des Systems festzulegen. In meiner Terminologie wird ein Netzwerk, das zwei Schichten verbindet, dadurch bestimmt, daß die Aktionsfunktion, $K(\Delta)$, für das elementare Netzwerk definiert wird. Das ganze Netzwerk wird dann einfach dadurch generiert, daß das elementare Netzwerk entlang aller Punkte der begrenzenden Schichten verschoben wird. In der nun folgenden Erörterung variabler Aktionsfunktionen halte ich weiter an einer genetisch programmierten Aktionsfunktion, sagen wir $K_0(\Delta)$, fest, ich betrachte diese Aktionsfunktion nun aber unter dem Einfluß gewisser Stör-Einwirkungen. Entsprechend meiner früheren Behauptung (Gleichung 14), daß wir uns diese Aktionsfunktion als aus zwei Teilen zusammengesetzt denken, und zwar aus einem strukturellen Teil $D(\Delta)$ und einem funktionalen Teil $T(\Delta)$, können wir nun die Störeinwirkung auf T, oder auf D, oder auf beide, betrachten. Auch wenn mir klar ist, daß die Annahme struktureller Variationen in Nervennetzen nicht allzu populär ist, möchte ich als mein erstes Beispiel die Ergebnisse gerade einer solchen strukturellen Störeinwirkung anbieten, die Sie vielleicht als eine mögliche frühe Phase in der Bildung solcher Netzwerke akzeptieren können.

Nehmen wir an, daß das einfachste aller möglichen elementaren Netzwerke genetisch programmiert wird (Bild 4).

Dieses Netz besteht aus präzise zwei Elementen, einem in der Schicht L_1, einem in der Schicht L_2. Da $2n = 2$, haben wir $n = 1$ und die Unsicherheit dieses Netzes als

$$H = 1 \text{ Bit/Netz}.$$

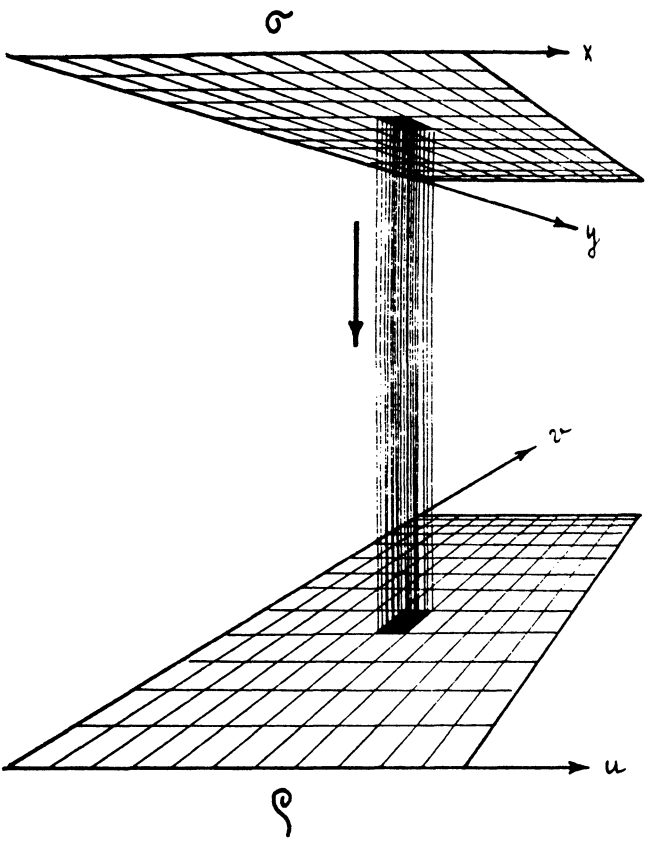


Bild 4 Geometrische Beziehungen in einer Aktionsfunktion, die durch eine ideale eins-zu-eins Abbildung definiert ist.

Das genetische Programm sagt: „Verknüpfe diese beiden Elemente!" Mathematisch gesprochen, läßt sich die Abbildungsfunktion $D(\Delta)$, die die Struktur des Netzes ausdrückt, durch die Diracsche Deltafunktion darstellen:

$$D(\Delta) = \delta(|\Delta|),$$

wobei

$$\delta(x - x_0) = \begin{cases} 0 \text{ for } x \neq x_0 \\ \infty \text{ for } x = x_0 \end{cases},$$

und

$$\int_{-\infty}^{-\infty} \delta(x - x_0)\, dx = 1.$$

Der Einfachheit halber wollen wir annehmen, daß die Transferfunktion $T(\Delta)$ lediglich eine Konstante ist:

$$T(\Delta) = a.$$

Damit ergibt sich die Aktionsfunktion einfach als

$$K(\Delta) = a\delta(|\Delta|),$$

und die Reaktion in der Schicht L_2 auf einen gegebenen Stimulus in L_1 ist entsprechend Gleichung (13):

$$\rho(\mathbf{r}_2) = a\lambda \int_{L_1} \delta(|\Delta|)\, \sigma(\mathbf{r}_1)\, dA = a\lambda\sigma(\mathbf{r}_2).$$

Mit anderen Worten, die Reaktion ist eine identische Abbildung des Stimulus mit aufgrund des Faktors $a\lambda$ modifizierter Amplitude, wie in diesem sehr einfachen Verknüpfungsschema zu erwarten.

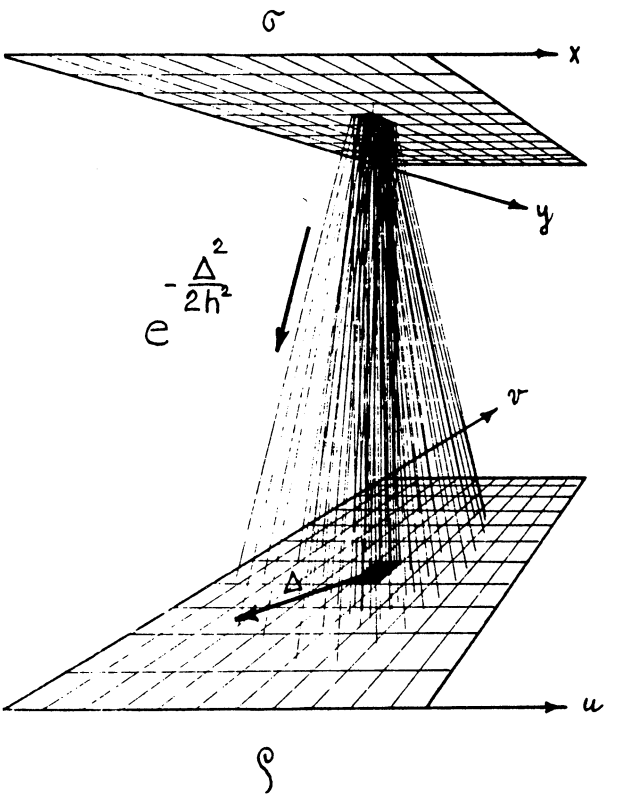


Bild 5 Geometrische Beziehungen in einer Aktionsfunktion, die sich aus einer zufallsabhängigen Störung einer idealen eins-zu-eins-Abbildung ergibt.

Nehmen wir nun an, daß es im Prozeß der Verwirklichung dieses Netzwerks unmöglich wird, daß die Fasern, die von L_1 ausgehen, direkten Kontakt mit L_2 herstellen, und daß sie aufgrund dazwischenliegender Gliazellen zufallsabhängige Umwege einschlagen müssen. Ein Bündel von Fasern, das bei **r** entsteht und bei $\mathbf{r}_2$ ankommen soll, wird nach der Gaußschen Verteilung gestreut. Die Abbildungsfunktion wird daher zu folgendem Ausdruck:

$$D(\Delta)\,\chi[\exp - \Delta^2/2h^2],$$

wobei h die Varianz dieser Verteilung repräsentiert (Bild 5).

Nehmen wir weiterhin an, daß die Transferfunktionen für inhibitorische und exzitatorische Verknüpfungen zwar Konstanten, für die zwei Arten $(+\,a_1)$ und $(-\,a_2)$ jedoch verschieden sind. Führen wir nun die entsprechenden Varianzen h_1 und h_2 ein, dann wird die Aktionsfunktion des elementaren Netzwerks zu

$$K(\Delta) = a_1 \exp[-\Delta^2/2h_1^2] - a_2 \exp[-\Delta^2/2h_2^2],$$

und die Stimulus-Reaktions-Relation kann mit Hilfe der Gleichung (13) ermittelt werden. Für einen einheitlichen Stimulus, $\sigma = \sigma_0$, ist die sich ergebende Reaktion gleichermaßen einheitlich:

$$\rho(r_2) = \sigma_0 \lambda(a_1 h_1^2 - a_2 h_2^2) = \text{konst.}, \tag{15}$$

und sie löst sich auf unter der Bedingung

$$a_1 h_1^2 = a_2 h_2^2. \tag{16}$$

Diese Bedingung ist der graphischen Abbildung der Gaußschen Aktionsfunktion in Bild 6 und auch in Bild 7 zugrundegelegt; die letztere zeigt die Reaktionsaktivität in der Schicht L_2, wenn das Stimulusmuster ein einheitlich beleuchtetes Quadrat ist. Das interessante Merkmal des zusammengesetzten Bildes in Bild 6 ist eine deutlich merkbare laterale Inhibition derjenigen Faserbündel, die auf die in L_2 lokalisierten Neuronen einwirken. Die Folge dieser Aktionsfunktion ist die Berechnung einer Kontur durch die Elemente von L_2. Diese Kontur tritt am deutlichsten dann hervor, wenn einheitliche Stimulierung keine Reaktion hervorruft, oder mit anderen Worten, wenn die durch Gleichung (16) ausgedrückte Bedingung erfüllt ist.

Da in unserem Programm nichts angenommen wurde, was eine dermaßen anspruchsvolle Bedingung festgelegt hätte, läßt sich als einziger Mechanismus für diese Leistung nur die Anpassung verantwortlich machen. Mit anderen Worten, die Aktionsfunktion, die sich bis jetzt entwickelt hat, muß weiteren Veränderungen unterworfen werden, wenn die idealen Verknüpfungsbedingungen, wie sie in Gleichung (16) ausgedrückt sind, nicht durch irgendwelche glücklichen Zufälle spontan hergestellt worden sind. Die Aussichten für einen solchen Glücksfall sind jedoch äußerst gering, und wir können erwarten, daß sich das Netzwerk nach seiner Bildung in einem Zustand befindet, in dem das Gleichheitszeichen der Gleichung (16) durch ein Ungleichheitszeichen ersetzt werden muß.

Es stellen sich nun die Fragen, was sich verändern wird, wie diese Veränderung bewerkstelligt wird, und was die Ursache dieser Veränderung ist. Wenn ich einige der Ausführungen im Verlauf dieser Tagung richtig verstanden habe, dann ist es mir nicht gestattet, die Struktur des Netzwerks, wie sie bisher entwickelt worden ist, zu verändern, denn nach der Herstellung der synaptischen Verknüp-

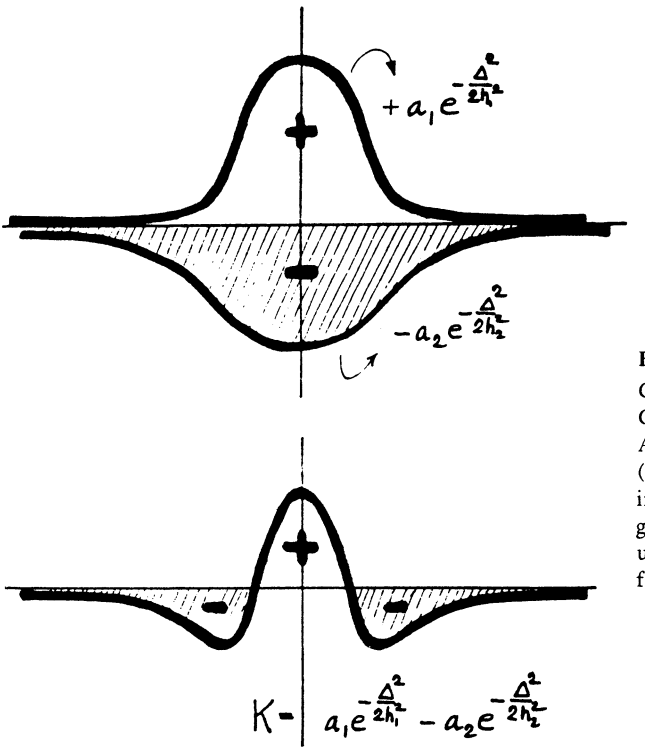


Bild 6
Graphische Darstellung einer Gaußschen (normal-)verteilten Aktionsfunktion.
(a) Exzitatorische und inhibitorische Verteilung getrennt. (b) Exzitatorische und inhibitorische Verteilungsfunktion verbunden.

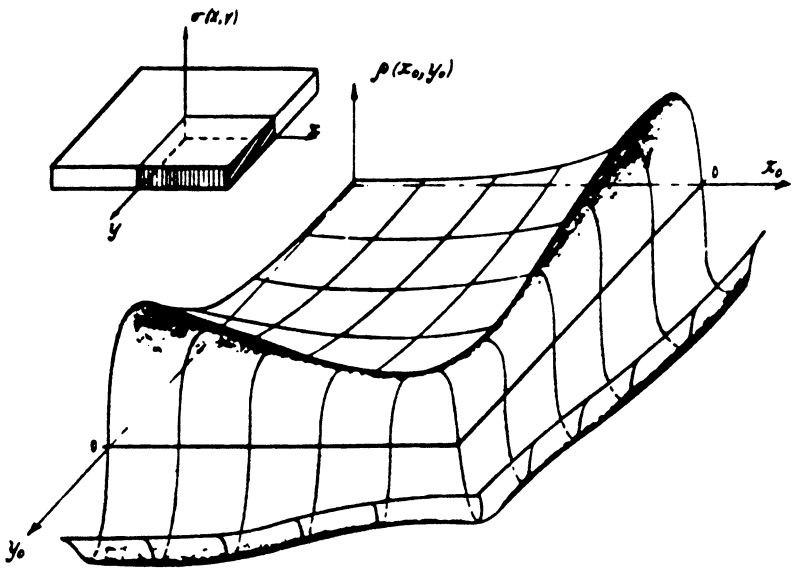


Bild 7 Reaktionsverteilung, hervorgerufen durch einen einheitlichen Stimulus, beschränkt auf ein Quadrat. Die Konturbestimmung ist die Folge einer Aktionsfunktion, die durch Überlagerung einer exzitatorischen und einer inhibitorischen Gaußschen Verteilungsfunktion erreicht wird.

fungen werden diese ganz starr aufrechterhalten. Der einzige mir verbleibende Ausweg besteht in einer Veränderung von $T(\Delta)$, der Transferfunktion, oder, wie ich bereits früher meinte, der operationalen Modalität der Knotenelemente dieses Netzwerks.

ROBERTS:

Warum ist es Ihnen verboten, die Struktur zu verändern?

von FOERSTER:

Ich folge lediglich den Vorschlägen, die hier schon gemacht wurden, meine ich, und die erlauben mir nicht, Synapsen herumzuschieben. Sie sind festgelegt.

SPERRY:

Das können Sie aber nicht machen, denn ich habe auf dieser Tagung bisher zumindest von seiten der Chemie und Mikrobiologie keinerlei deutliche Einwände gehört, als wir uns mit der Vorstellung befaßten, daß Synapsen strukturell verändert werden könnten.

von FOERSTER:

Ja, aber in meiner Terminologie handelt es sich dabei um funktionale Veränderungen der Transferfunktion, die durch irgendwelche submikroskopischen Veränderungen der synaptischen Struktur verursacht werden. Ich weiß nicht, ob man sie unter dem Mikroskop sehen kann. Ich möchte nur noch anfügen, daß ich in der Tat sehr glücklich wäre, wenn ich die Struktur des Netzwerks verändern könnte. Was ich aber zu sagen versuche, ist, daß ich auch dann, wenn keine Strukturveränderungen in meinen Netzwerken zugelassen werden, immer noch in der Lage bin, einen Rechner zu bauen, der seine eigene innere Organisation verändert, denn ich kann dann die Transferfunktionen meiner elementaren Bestandteile verändern. Ich denke, daß Sir John eben das gezeigt hat, als er auf die Variation der synaptischen Übertragungsleistung hinweis.

Ich komme also wieder auf meine Aktionsfunktion zurück und schlage vor, den strukturellen Teil $D(\Delta)$, der als eine zufallsbedingte Verteilung der verknüpften Fasern bestimmt ist, nicht zu verändern, die lokalen Transferfunktionen aber, $T(\Delta)$, bestimmten von außen kommenden Störeinwirkungen zu unterwerfen. Die von mir früher eingeführte Transferfunktion war von der allereinfachsten Art. Sie bestand aus Exzitation bzw. Inhibition. Nehmen wir nun an, daß nach Herstellung der Kontakte diese beiden Konstanten bestimmte Ausgangswerte haben, sagen wir $(+ a_{10})$ und $(- a_{20})$. Die entscheidende Bedingung für optimale Konturextraktion (Gleichung 16) ist jedoch zum Anfang noch nicht erfüllt:

$$a_{10} h_1^2 \neq a_{20} h_2^2 .$$

Da die Varianzen der inhibitorischen und exzitatorischen Verknüpfungen in dieser Gleichung strukturelle Parameter (und nach meinen Regeln tabu) sind, muß ich irgendetwas auf entweder a_{10} oder a_{20} einwirken lassen, um die gewünschte Gleichheit herzustellen. Unter den vielen Möglichkeiten, die sich von einem rein hypothetischen Gesichtspunkt her anbieten, möchte ich eine vorschlagen, die zumindest für mich einen gewissen Anschein von Plausibilität hat. Sei zum Beispiel die linke Seite der Ungleichung größer als die rechte Seite. Ich werde nun annehmen, daß zumindest eine dieser Transferfunktionen, in diesem Falle a_2, das heißt die inhibitorische Transferfunktion, ihre Leistung aufgrund der Gesamtakti-

vität des reagierenden Netzwerks verstärkt. Mit anderen Worten, ich mache diese Rückkopplungsschleife nicht zu einer lokalen Angelegenheit, sondern vielmehr zu einer Sache, an der die Aktivität ganzer Zellkomplexe einschließlich der sie umgebenden Gewebe mitwirkt. Exakt formuliert ergibt sich

$$\frac{\partial a_2}{\partial t} = cR,$$

wobei R für die Gesamtreaktion der Schicht L_2 steht:

$$R = \int_{L_2} \rho(\mathbf{r}_2)\, dA.$$

Um nun ein zumindest grobes Bild dessen zu bekommen, was unter diesen Bedingungen geschieht, wollen wir ein trübes Universum für unser System annehmen, in dem nur gelegentlich Objekte auftreten, die von den Elementen in L_1 „gesehen" werden, das ansonsten aber einheitlich erleuchtet ist. In diesem Falle läßt sich R sofort mit Hilfe der Gleichung (15) berechnen:

$$R = (a_{10}h_1^2 - a_2h_2^2) \cdot S(t),$$

wobei S(t) die Gesamtstimulierung ist, die auf die Schicht L_1 als eine Funktion der Zeit einwirkt. Setzen wir dies in die Gleichung (17) ein, ergibt sich in a_2 eine Differentialgleichung:

$$\frac{\partial a_2}{\partial t} = c\lambda(a_{10}h_1^2 - a_2h_2^2) \cdot S(t),$$

mit der Anfangsbedingung

$$t = 0 \ldots a_2 = a_{20}$$

und ihrer Lösung:

$$a_2h_2^2 = a_{10}h_1^2 - (a_{10}h_1^2 - a_{20}h_2^2) \exp - h_2^2 c\lambda \int_0^t S(t)\, dt.$$

Der erste Ausdruck repräsentiert die durch Gleichung (16) ausgedrückte gewünschte Bedingung, während der zweite aufgrund des abnehmenden Exponentialausdrucks letztlich verschwinden muß, wie immer die Geschichte des Stimulus $\int_0^t S dt$ ausgesehen haben mag.

Auch wenn dies zugegebenerweise ein fast triviales Beispiel für ein adaptives Netzwerk ist, hat es mir doch ohne allzu große Akrobatik gestattet, meine Auffassung deutlich zu machen, daß sich ein Rechner, der anfangs als idealer Wiederholer angesehen wurde, wenn er der Härte einer wirklichen Welt ausgesetzt wird, zu einem perfekten Konturendetektor wandelt. Dies ist um so überraschender, wenn man sich einen Augenblick überlegt, daß diese Maschine von beachtlicher Kompliziertheit durch eine genetische Anweisung entstanden ist, die, wie sie sich vielleicht erinnern, nur ein einziges Bit an Information erfordert hat!

Ich könnte mich nun nach diesem Beispiel zurückziehen, gäbe es da nicht einen weiteren Einwand, der mir im Kopf herumspukt: ich könnte nämlich erneut beschuldigt werden, lediglich die Geschichte eines adaptiven Filtermechanismus dargestellt zu haben, ohne das tiefe Problem des Gedächtnisses auch nur berührt zu haben.

Aus diesem Dilemma werde ich glücklicherweise durch eine hübsche, wenn auch nicht unmittelbar einsichtige Eigenschaft der Gaußschen Aktionsfunktion gerettet. Es zeigt sich nämlich, daß dann, wenn die Bedingung der Gleichung (16) erfüllt ist, dieses Verknüpfungsschema in der Nachbarschaft der Punkte $\mathbf{r}_2$ entlang der Kontur der Reaktionsschicht L_2 eine mittlere Reaktionsdichte $\rho(\mathbf{r}_2)$ erzeugt, die der Krümmung der Kontur proportional ist. Nehmen wir $\underline{R}(\mathbf{r}_2)$ als Radius der Krümmung der Kontur am Punkt $\mathbf{r}_2$ an, und sei k eine Konstante, dann ergibt sich

$$\rho(\mathbf{r}_2) = k/\underline{R}(\mathbf{r}_2).$$

In der ersten Annäherung ist die totale Reaktion R die über die gesamte Kontur integrierte Aktivität

$$R = \int \rho(\mathbf{r}_2)\, ds = k \int \frac{ds}{\underline{R}(\mathbf{r}_2)},$$

wobei ds ein Zeilenelement entlang der Kontur darstellt. Aber

$$ds = \underline{R}(\mathbf{r}_2)\, d\psi,$$

wobei ψ den Polwinkel vom Zentrum der Krümmung repräsentiert. Daher

$$R = k \int_{\psi_1}^{\psi_2} \frac{\underline{R}(\mathbf{r}_2)\, d\psi}{\underline{R}(\mathbf{r}_2)} = k \int_{\psi_1}^{\psi_2} d\psi = \text{größeninvariant}.$$

Mit anderen Worten, die einzige Funktion, die die Information über die Größe des Objektes transportiert, nämlich $\underline{R}$, hebt sich weg, und die sich ergebende Reaktion R ist mit Bezug auf die Größe verschiedener Objekte invariant und variiert lediglich mit ihrer Gestalt.

Stellen wir uns nun vor, daß unser System in Kontakt steht mit einem Universum, das die eigentümliche Eigenschaft hat, durch eine festgelegte Zahl von Objekten jeweils unterschiedlicher Größe, aber ähnlicher Gestalt bevölkert zu sein, die sich frei herumbewegen, so daß das begrenzte „visuelle Feld" unseres Systems in jedem Augenblick nur eine bestimmte Anzahl dieser Objekte wahrnimmt. Da jedes gesehene Objekt die gleiche Reaktion R_0 hervorrufen wird, rufen Q Objekte im visuellen Feld eine Gesamtreaktion $Q \cdot R_0$ hervor. Mit anderen Worten, unser System ist in der Lage zu zählen! Da nichts dergleichen zu den Eigenschaften des Systems gehörte, als es auf die Welt kam — wir haben es immer noch mit nur einem Bit genetischer Information zu tun —, muß das System seine mathematischen Fähigkeiten im Laufe seiner Interaktionen mit seiner Umwelt erworben haben. Man kann außerdem eine große Vielfalt von Lehrmethoden entwickeln, die den Prozeß beschleunigen, durch den das System sein Wissen über Zahlen erwirbt. Hat es schließlich diese Aufgabe gemeistert, kann ein Lehrer, der die inneren Vorgänge unseres Systems überhaupt nicht kennt, über die Lokalisierung seines Gedächtnisses spekulieren. Wir wissen nun natürlich, daß dieses Gedächtnis *überall*

ist, daß es in der Struktur des Verknüpfungsschemas und in den operationalen Modalitäten aller Knotenpunkte dieses Netzwerks verwirklicht ist.

Ich möchte nun meine Ausführungen beschließen, indem ich wiederum einige Zahlen gebe. Lassen sie mich zunächst unseren Lehrer hernehmen, der die einfache evolutionäre Geschichte unseres Systems nicht kennt, der seine Anatomie studieren will, und der also ein außergewöhnlich kompliziertes Netzwerk vorfindet, das aus n elementaren Netzen besteht, die alle verschieden sind. Mit guter Statistik kann er seine Daten dadurch erfassen, daß er zeigt, daß die Verknüpfungen aller elementaren Netze einer Gaußschen Verteilung unterliegen, wobei jedes Netzwerk im Durchschnitt etwa m verknüpfte Elemente enthält. Die Unsicherheit H_e für das einzelne elementare Netzwerk ist annähernd[2)]

$$H_e = \log_2 \sqrt{m},$$

und für das Netzwerk, das aus N Elementen besteht,

$$H = N \log_2 \sqrt{m}.$$

Benutze ich nun die Zahlen der Tabelle 4, so wie ich schon früher vorgeschlagen habe, dann ist die Unsicherheit dieses Systems, das aus nur einem Netz (C = 1) besteht:

$$H = 3 \cdot 10^4 \log_2 \sqrt{300} = 1{,}2 \cdot 10^5 \text{ Bits}.$$

Der Anatom kann dies dem genetischen Programm zuschreiben und dessen Informationsgehalt auf etwa 100000 Bits schätzen. Wir wissen jedoch, daß die ursprünglichen Bestimmungen mit Hilfe nur eines einzigen Bits geschrieben wurden. Woher kommt also all diese Information? Die Antwort ist natürlich: von all den „Störungen", die auf das Netzwerk eingewirkt haben, als es seine schwachen Versuche machte, den genetischen Befehl auszuführen. Es ist, wie in vielen anderen Fällen, auch hier so, daß es die Störeinwirkungen sind, die eine zunächst sehr kümmerliche Struktur bereichern.

Wir wollen uns nun der Arbeitsweise und Leistungsfähigkeit unseres Systems zuwenden. Gibt es N unabhängige binäre Elemente in seiner „sensorischen" Schicht L_1, die jeweils nach einem Zeitintervall von t_0 Sekunden für neue Sinneswahrnehmungen bereit sind, dann ist die Aufnahmefähigkeit für Eingabeinformation des Systems

$$H_{in} = N/t_0 \text{ Bits/Sekunde}.$$

Seine Ausgabeleistung hängt natürlich von den Anforderungen der Systemumwelt ab. Nehmen wir an, daß die Anzahl der Objekte, die von unserem Lebewesen gesehen werden, entsprechend der Normalverteilung variiert, mit einer Varianz von, sagen wir, M Objekten. Das System berichtet damit in Intervallen von t_0 Sekunden über den Zustand seines visuellen Universums. Seine informationelle Output-Leistung ist daher:

$$H_{out} = \frac{1}{t_0} \log_2 \sqrt{M},$$

---

2) Aufgrund der Einschränkungen in einer normalen Wahrscheinlichkeitsverteilung ist die Unsicherheit nicht $\log_2 m$, sondern nur annähernd $1/2 \log_2 m$.

und das Verhältnis zwischen Output und Input ist:

$$H_{out}/H_{in} = (\log_2 \sqrt{M})/N.$$

M ist jedoch immer kleiner als N, denn man kann nicht mehr Gegenstände sehen, als es Zapfen und Stäbchen gibt. Im optimistischesten Fall ist M = N und das optimale Output-Input-Verhältnis ist

$$\left(\frac{H_{out}}{H_{in}}\right)_{opt} = \frac{\log_2 N}{2N}.$$

Bei N = 30000 wird dieses Verhältnis wiederum zu:

$$\left(\frac{H_{out}}{H_{in}}\right)_{opt} = 0{,}00075.$$

Mit anderen Worten, das System reduziert die Unsicherheit seiner Umwelt beträchtlich, da es die Fähigkeit besitzt, zu klassifizieren und zu abstrahieren. In diesem Falle bedeutet dies die Identifikation von einzelnen Objekten.

Ich habe bis jetzt lediglich die Wunder herausgestellt, die gewirkt werden können, wenn Unordnung in ein geordnetes genetisches Programm eingeführt wird. Man kann sich nun fragen: Wo ist die innere Organisation, die in meinem System durch Interaktionen mit seiner Umwelt entsteht, wie ich dies mit meinen früheren Behauptungen betont habe? Es ist klar, daß die Folgen derartiger Interaktionen erst dann erklärt werden können, wenn das System zur Interaktion überhaupt fähig ist; dies ist aber erst möglich, sobald das Netzwerk sich selbst gebildet hat. Die einzige Quantität, die sich danach verändert, ist die inhibitorische Transferfunktion a_2, die ich der Einfachheit halber durch ein „a" ohne Index benennen will. Sei a_{min} das kleinste in der Beobachtung dieser Transferfunktion erkennbare Intervall, dann wird n, die Anzahl der Zustände dieses „Universums", durch den Maximumwert a_{max} bestimmt, den diese Funktion annehmen kann:

$$n = a_{max}/a_{min}.$$

Daraus ergibt sich H_{max}, die maximale Unsicherheit:

$$H_{max} = \log_2 n.$$

Durch

$$n_0 = a_0/a_{max} \ll n$$

ergibt sich seine anfängliche Unsicherheit als

$$H_0 = \log_2 (n - n_0),$$

das heißt, als Logarithmus der Anzahl noch verfügbarer Zustände. Folglich ist sein Anfangsmaß der Organisation:

$$\Omega_0 = 1 - \frac{H_0}{H_m} = \frac{n_0/n}{\log_2 n}.$$

Wenn das System nach dem Interaktionskontakt in seinen Reifezustand übergeht, sagen wir nach n_1 unterscheidbaren Zuständen, ergibt sich der Endzustand seiner Organisation als:

$$\Omega_\infty = 1 - \frac{\log_2 n_1}{\log_2 n} = \frac{\log_2 (n/n_1)}{\log_2 n} .$$

Der Vergleich zwischen dem Ausgangs- und dem Endzustand seiner Organisation zeigt, daß das Maß des Endzustandes der Organisation größer ist als das des Ausgangszustandes der Organisation, wenn

$$\log_2 (n/n_1) > (n_0/n),$$

oder wenn näherungsweise

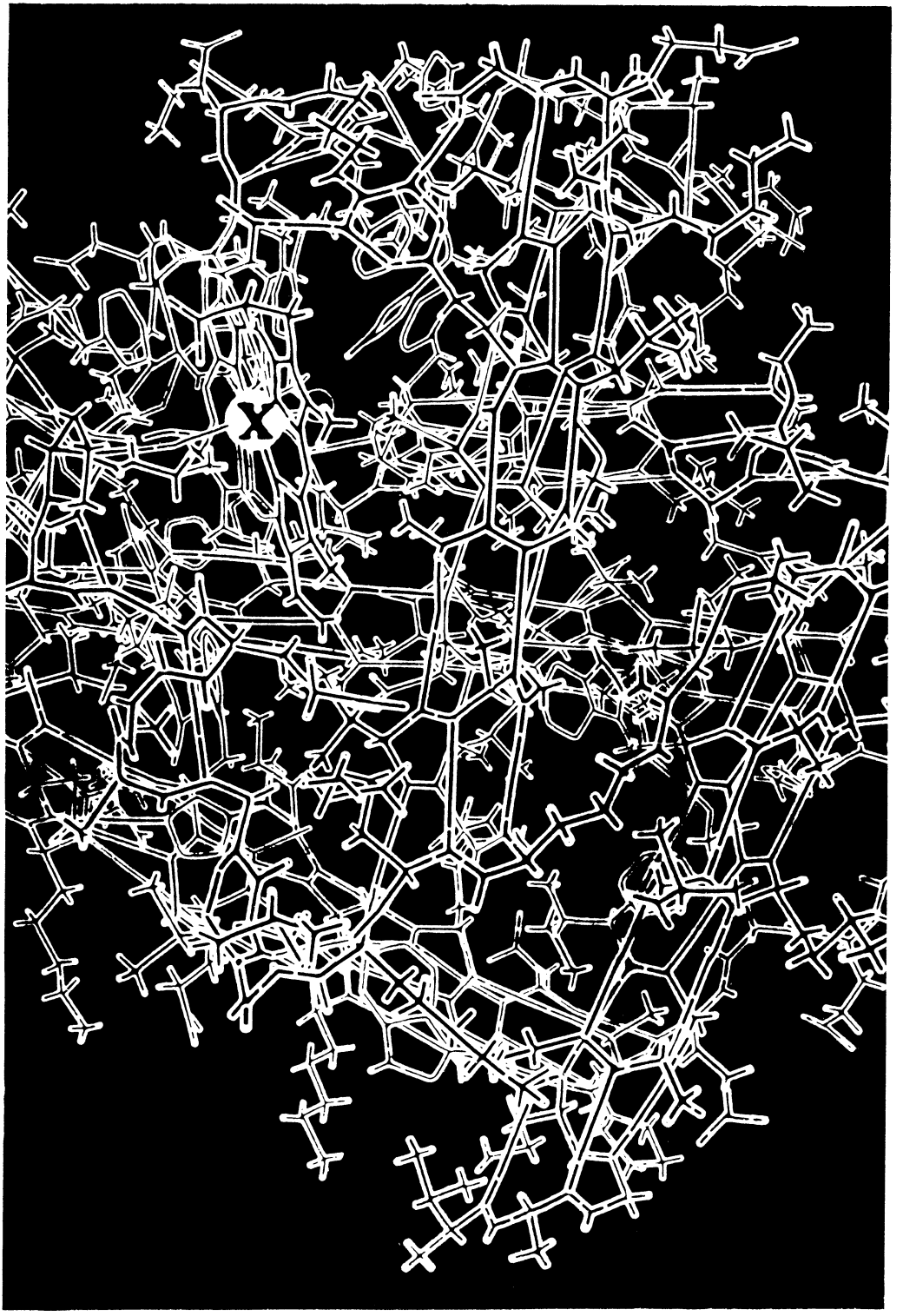
$$n > n_0 + n_1 .$$

Dies ist jedoch immer der Fall. Wenn also die Schwankung um den endgültigen Gleichgewichtszustand äußerst gering ist, sagen wir $\bar{n}_1 = 1 + E$, dann ergibt sich

$$\Omega_\infty \approx 1 - [\epsilon/\log_2 n] \approx 1,$$

und das System ist in fast vollkommener Ordnung.

Ich hoffe, daß diese Bemerkungen die Möglichkeit deutlich machen, lernende und sich erinnernde Systeme als Rechner aufzufassen, die aufgrund ihrer Interaktionen mit ihrer Umwelt ihre eigenen operationalen Modalitäten verändern. Ihre Operationen verändern sich in solcher Weise, daß sie immer mehr Unsicherheit der Umwelt beseitigen, bis der Output der Systeme sie in Gleichgewicht mit ihren Universen erhält.



Molekular-Ethologie: ein unbescheidener Versuch semantischer Klärung

I Einführung

Die molekulare Genetik ist ein Beispiel für den erfolgreichen Brückenschlag zwischen einer Phänomenologie unmittelbar erfahrener makroskopischer Dinge (einer Taxonomie von Arten, Variationen innerhalb einer Art usw.) und der Struktur und Funktion einiger weniger mikroskopischer Elementareinheiten (in diesem Falle einer spezifischen Menge organischer Makromoleküle), deren Eigenschaften aus anderen, eigenständigen Beobachtungen abgeleitet worden sind. Ein wichtiger Schritt in diesem Brückenbau ist die Erkenntnis, daß diese elementaren Einheiten nicht notwendig die alleinigen Konstituenten der an Dingen beobachtbaren makroskopischen Merkmale sind, sondern vielmehr Determinanten der Synthese von Einheiten, die die makroskopischen Entitäten aufbauen. Gleichermaßen hilfreich ist die Metapher, die die elementaren Einheiten als „Programm" und die synthetisierten Bestandteile in ihrer makroskopischen Erscheinung als Ergebnis einer „Errechnung" auffaßt, welche durch das entsprechende Programm eingeleitet und gesteuert wird. Die Gene, die blaue Augen festlegen, sind nicht selbst blaue Augen, in blauen Augen jedoch finden sich Kopien der Gene, die die Entwicklung blauer Augen bestimmen.

Angefeuert durch den Erfolg der molekularen Genetik, fühlt man sich nun versucht, nach der Brücke zu suchen, die eine andere Menge makroskopischer Phänomene, nämlich das Verhalten von Lebewesen, mit der Struktur und der Funktion einiger weniger mikroskopischer Elementarbausteine verbindet, höchstwahrscheinlich den gleichen, die für Gestalt und Organisation des lebenden Organismus verantwortlich sind. Die „molekulare Ethologie" ist jedoch bis jetzt noch nicht von Erfolg gekrönt, und es scheint empfehlenswert, die Ursachen dafür zu klären.

Eine Ursache scheint darin zu liegen, daß die kognitiven Fähigkeiten des Menschen, Formen und Gestalten zu unterscheiden und zu identifizieren, im Vergleich zu seinen Fähigkeiten der Unterscheidung und Identifizierung von Veränderungen und Bewegungen ungleich größer sind. Es gibt in der Tat einen Unterschied zwischen diesen beiden kognitiven Prozessen, der sich auch in der semantischen Struktur jener Sprachelemente niederschlägt, die diese beiden Erscheinungsarten bezeichnen; wir verwenden nämlich verschiedene Substantive für Gegenstände, die sich nach Form und Gestalt unterscheiden, und Verben für Veränderungen und Bewegung.

Der strukturelle Unterschied zwischen Substantiven (cl_i^k) und Verben (v_i) wird deutlich, wenn dafür lexikalische Definitionen angefertigt werden. Im allgemeinen bezeichnet ein Substantiv eine Klasse (cl^1) von Objekten. Wird ein Substantiv definiert, dann wird es als Element einer umfassenderen Klasse (cl^2) aus-

gewiesen, die ihrerseits ebenso durch ein Substantiv bezeichnet wird, welches seinerseits durch Definition als Element einer wiederum umfassenderen Klasse (cl^3) ausgewiesen wird usw. [Fasan → Vogel → Tier → Organismus → Ding]. Dies entspricht dem folgenden Schema der Repräsentation des Definitionsparadigmas für Substantive:

$$cl^n = \{ cl^{n-1}_{i_{n-1}} \{ cl^{n-2}_{i_{n-2}} \{ \ldots \{ cl^{m}_{i_m} \}\}\}\}, \qquad (1)$$

wobei die Notation $\{\epsilon_i\}$ für eine Klasse von Elementen ϵ_i (i = 1, 2, ..., p) steht, und die subskribierten Subskripte dazu dienen, diese Subskripte mit den entsprechenden Superskripten zu verknüpfen. Die höchste Ordnung n in dieser Hierarchie von Klassen wird stets durch einen einzigen undefinierten Begriff, wie „Ding", „Entität", „Akt", etc. repräsentiert, der sich stets auf grundlegende Vorstellungen der allgemeinen Wahrnehmungsfähigkeit schlechthin bezieht. Eine graphische Darstellung der hierarchischen Ordnung von Substantiven bietet Bild 1; eine detaillierte Erörterung der Eigenschaften solcher (umgekehrter) „Substantivketten-Bäume" ist an anderer Stelle verfügbar (Weston 1964; von Foerster 1967a).

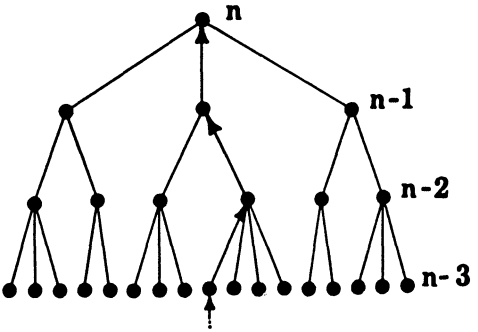


Bild 1 Aufsteigende hierarchische Definitionsstruktur für Substantive. (Substantive: Knoten; Pfeilspitzen: Definiens; Pfeilenden: Definiendum.)

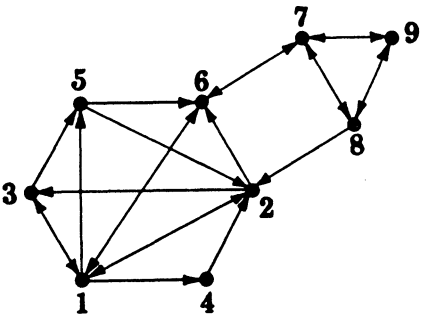


Bild 2 Geschlossene heterarchische Definitionsstruktur für Verben. (Verben: Knoten; Pfeilspitzen: Definiens; Pfeilenden: Definiendum.)

Ein Verb (v_i) bezeichnet im wesentlichen eine Handlung, seine Definition wird gewöhnlich durch eine Menge von Synonymen $\{v_j\}$, d.h. durch die Vereinigungs- oder die Schnittmenge der Bedeutungen von Verben gegeben, die ähnliche Handlungen bezeichnen: [schlagen → {klopfen, prügeln, pulsieren} → {(schlagen, prügeln, ...) (hauen, dreschen, kämpfen, ringen, meistern, züchtigen, vergelten, ... balzen) (klopfen, pulsieren, schwingen, stampfen, schütteln, stoßen ...)} → etc.]

$$v_i = \{v_j\} \vee \sum v_k \vee \prod v_e \qquad (2)$$

Eine graphische Darstellung dieser im Grunde geschlossenen heterarchischen Struktur bietet Bild 2; die entsprechende Abbildung in Form endlicher Matrizen ist an anderer Stelle erörtert worden (von Foerster 1966).

Der wesentliche Unterschied zwischen den kognitiven Prozessen, die die Identifikation von Formen, und jenen, die die Identifikation von Veränderungen von Formen ermöglichen, spiegelt sich nicht nur in den völlig unterschiedlichen Formalismen, die für die Abbildung der verschiedenen Definitionsstrukturen von Substantiven [Gleichung (1)] und Verben [Gleichung (2)] benötigt werden, sondern auch in der Tatsache, daß die Menge der Invarianten, die eine Gestalt in verschiedenen Transformationen identifizieren, durch einen einzigen *deduktiven* Algorithmus berechnet werden kann (Pitts/McCulloch 1947), während die Identifizierung schon der elementarsten Verhaltenskonzepte *induktive* Algorithmen erfordert, die nur durch ständigen Vergleich der gegenwärtigen Zustände des betrachteten Systems mit seinen vorausgegangenen Zuständen berechnet werden können (von Foerster u. a. 1968).

Diese kognitiven Handikaps machen es dem Ethologen ungleich schwerer als seinem Kollegen, dem Genetiker, eine Phänomenologie für seinen Gegenstandsbereich zu entwickeln. Den Werkzeugen zur Beschreibung seiner Phänomene geht nicht nur die schöne Isomorphie ab, die zwischen der hierarchischen Struktur aller Taxonomien und den Definitionen von Substantiven gilt, die diese bezeichnen, er kann vor allem leicht in eine semantische Falle gehen, indem er einer begrifflich isolierbaren Funktion einen entsprechend isolierbaren Mechanismus unterstellt, der diese Funktion generieren soll. Diese Verführung scheint mir ganz besonders stark zu sein dann, wenn unser Wortschatz eine Vielfalt von begrifflich trennbaren höheren geistigen Fähigkeiten anbietet, wie zum Beispiel „lernen", „merken", „sich erinnern", „wahrnehmen" „vorhersagen" usw., und wenn der Versuch gemacht wird, in den verschiedenen Teilen unseres Gehirns die Mechanismen zu identifizieren und zu lokalisieren, die jeweils lernen, merken, wahrnehmen, sich erinnern, vorhersagen usw. Die Aussichtslosigkeit der Suche nach Mechanismen, die ausschließlich eine dieser Funktionen erfüllen, hat keinen physiologischen Grund, wie etwa „die große Komplexität des Gehirns", „die Schwierigkeit zu messen" usw. Diese Aussichtslosigkeit hat eine rein semantische Grundlage. Das in Isolation untersuchte „Gedächtnis" etwa wird so reduziert auf „Aufzeichnungen", „Lernen" auf „Veränderung", „Wahrnehmen" auf „Input" usw. Mit anderen Worten, durch die Abtrennung dieser Funktionen von der Totalität der kognitiven Prozesse hat man das ursprüngliche Problem verlassen und sucht nun nach Mechanismen, die völlig andere Funktionen erfüllen, und die gewissen Prozessen, die der Aufrechterhaltung des Organismus als einer funktionierenden Einheit dienen, ähnlich sein mögen oder auch nicht (Maturana 1970a).

Man betrachte einmal zwei denkbare Definitionen des Gedächtnisses:

(A) die potentielle Vergegenwärtigung der vergangenen Erfahrungen eines Organismus;

(B) die beobachtete Veränderung der Reaktion eines Organismus auf ähnliche Abfolgen von Ereignissen.

Während die Definition A eine Fähigkeit (Gedächtnis $_A$) für einen Organismus postuliert, dessen innere Erfahrung einem außenstehenden Beobachter nicht zugänglich ist, postuliert Definition B die gleiche Fähigkeit (Gedächtnis $_A$) als nur im Beobachter wirksam – denn andernfalls hätte dieser den Begriff der „Veränderung" nicht entwickeln können –, ignoriert jedoch diese Fähigkeit im beobachteten Organismus, denn ein Beobachter kann an den inneren Erfahrungen des Organismus „prinzipiell" nicht teilhaben. Daraus ergibt sich Definition B.

Nach herrschender Meinung gehorcht am ehesten die Definition B den Grundregeln der „wissenschaftlichen Methode", so als ob es unmöglich wäre, Selbstreferenz, Selbstbeschreibung und Selbsterklärung wissenschaftlich zu behandeln, d.h. geschlossene logische Systeme, die den Bezugsgegenstand (Referenten) in die Bezugnahme (Referenz), den Beschreiber in die Beschreibung, und die Axiome in die Erklärung einbeziehen.

Eine solche Auffassung ist unbegründet. Derartige logische Systeme sind nicht nur eingehend untersucht worden (z.B. von Günther 1967; Löfgren 1968), auch Neurophysiologen (Maturana u.a. 1968), Experimentalpsychologen (Konorski 1962) und andere (Pask 1968; von Foerster 1969) sind zu solchen Konzepten vorgestoßen.

Diese Vorbemerkungen zeigen, daß die Erforschung der Mechanismen des Denkens zwei Arten von Problemen zu lösen hat, und daß nur eine Art davon zur Physiologie bzw. zur Physik gehört; die andere gehört zur Semantik. Es bedarf folglich einer erneuten Prüfung gewisser gängiger Begriffe des Lernens und des Gedächtnisses hinsichtlich der Kategorie, zu der sie gehören, sowie der Entwicklung eines theoretischen Rahmens, in dem diese Begriffe ihren angemessenen Ort haben.

Der nächste Abschnitt, „Theorie", prüft und definiert Begriffe, die im Rahmen eines einheitlichen mathematischen Formalismus mit Lernen und Gedächtnis verknüpft werden. Im Abschnitt III werden verschiedene Modelle der Interaktion von Molekülen mit funktionalen Einheiten höherer Organisation erörtert.

II Theorie

A Allgemeine Vorbemerkungen

Da wir noch nicht über eine umfassende Theorie des Verhaltens verfügen, haben wir auch keine Theorie des Lernens und folglich auch keine Theorie des Gedächtnisses. Nichtsdestoweniger gibt es heute ein großes Spektrum von Begriffssystemen, das von naivsten Interpretationen des Lernens bis zu subtilsten Erklärungsansätzen dieses Phänomens reicht. Am naiven Pol wird „Lernen" als eine Veränderung der Auftretenshäufigkeit der Handlungen eines Organismus interpretiert. Diese Handlungen sind im vorhinein zweifach festgelegt: einmal durch die Fähigkeit eines Experimentators, Handlungen überhaupt einzugrenzen, zum anderen durch sein Wertsystem, das Handlungen als „Treffer" oder als „Fehler" klassifiziert. Veränderungen werden durch die Manipulation des Organismus mit Hilfe von Elektroschocks, Nahrungsangeboten usw., oder in noch drastischerer Weise durch Verstümmelung oder sogar Entfernung bestimmter Organe des Organismus herbeigeführt. „Lehren" ist in dieser Perspektive die Anwendung von „Verstärkungen", die die bei anderer Gelegenheit beobachteten Veränderungen bewirken.

Am subtilen Pol wird Lernen als ein Prozeß der Entwicklung von Algorithmen angesehen, mit denen Kategorien von Problemen zunehmender Komplexität gelöst werden (Pask 1968), oder als ein Prozeß der Entwicklung von Relationsbereichen zwischen dem Organismus und der Außenwelt, von Relationen zwischen diesen Bereichen usw. (Maturana 1970a). Lehren ist in dieser Perspektive die Erleichterung derartiger evolutionärer Prozesse.

In einem direkten Verhältnis zum Grad der begrifflichen Verfeinerung aller dieser Ansätze steht ihre mathematische Primitivität: die begrifflich simplen Theorien verschleiern ihre Simplizität durch ein Gewölk mathematischen Aufwands; die differenzierten Theorien erweisen sich als völlig außerstande, ihre Tiefe zu vermitteln, da ihnen ein strenges, formales Instrumentarium abgeht. Unter den vielen Gründen für diese unbefriedigende Sachlage scheint einer besonders hervorzustechen: die außerordentlichen Schwierigkeiten, auf die jeder Versuch sofort stößt, mathematische Modelle zu entwickeln, die unserem erkenntnistheoretischen Verständnis angemessen sind. Vielleicht bedarf es des universalen Geistes eines John von Neumann, die adäquaten Werkzeuge dafür zu entwickeln. Da sie nun aber eben fehlen, wollen wir uns einmal in der mathematischen Werkstatt umsehen und prüfen, was zur Verfügung steht und unseren konkreten Zielen am besten dienen kann.

Ich habe in dieser Arbeit die Theorie der „Maschinen mit endlich vielen Zuständen" ausgewählt, um damit Reichweite und Grenzen bestimmter Konzepte in Theorien des Gedächtnisses, des Lernens und des Verhaltens zu demonstrieren, und dies vor allem aus zwei Gründen. Der erste Grund liegt darin, daß diese Theorie den direktesten Weg bietet, die externen Variablen eines Systems wie zum Beispiel Stimulus, Reaktion, Input, Output, Ursache, Wirkung usw. mit Zuständen und Operationen zu verknüpfen, die im Inneren des Systems angesiedelt sind. Da das zentrale Problem eines Buches über „molekulare Mechanismen des Gedächtnisses und des Lernens" darin bestehen muß, eine Verknüpfung zwischen solchen inneren Mechanismen und ihren Manifestationen im beobachtbaren Verhalten herzustellen, scheint die „Maschine mit endlich vielen Zuständen" ein nützliches Modell für diese Aufgabe zu sein.

Der zweite Grund für diese Wahl ist, daß die Interpretationen dieses Formalismus völlig offen bleiben können, daß sie das Lebewesen als Ganzes, Zellgruppen innerhalb des Lebewesens, einzelne Zellen und ihre operationalen Modalitäten, z.B. das einzelne Neuron, subzelluläre Bestandteile, und schließlich die molekularen Bausteine dieser Bestandteile betreffen können.

Mit der Bitte um Nachsicht an alle jene Leser, die an eine ausführlichere und strengere Behandlung dieses Themas gewöhnt sind, möchte ich die wichtigsten Teile dieser Theorie kurz skizzieren, um all jenen, die mit diesem Formalismus vielleicht nicht voll vertraut sind, die Mühe des Studiums weiterer Literatur zu ersparen (Ashby 1956; Ashby 1962; Gill 1962).

B Maschinen mit endlich vielen Zuständen

1 Deterministische Maschinen

Die Theorie der Maschinen mit endlich vielen Zuständen ist im wesentlichen die Theorie des Rechnens. Sie postuliert zwei endliche Mengen externer Zustände, die „Inputzustände" und „Outputzustände" genannt werden, eine endliche Menge „interner Zustände", und zwei explizit definierte Operationen (Berechnungen), die die gleichzeitigen sowie die historischen Relationen zwischen diesen Zuständen determinieren.[1])

[1]) Auch wenn die Interpretation von Zuständen und Operationen mit Bezug auf Observable völlig offengelassen wird, ist Vorsicht angebracht, sollen diese als mathematische Modelle etwa des Verhaltens eines lebenden Organismus dienen. Eine spezifische physikalische raumzeitliche Konfiguration, die von

Seien $\mathfrak{x}_i$ ($i = 1, 2, \ldots, n_x$) die n_x Rezeptoren für die Inputs x_i, deren jeder eine endliche Anzahl, $v_i > 0$, verschiedener Werte annehmen kann. Die Anzahl unterscheidbarer Inputzustände ist dann:

$$X = \prod_{i=1}^{n_x} v_i. \tag{3}$$

Ein bestimmter Inputzustand $x(t)$ zur Zeit t (oder kurz x) besteht dann in der Identifizierung der Werte x_i in allen n_x Inputrezeptoren $\mathfrak{x}_i$ in diesem „Moment":

$$x(t) \equiv x = \{x_i\}. \tag{4}$$

Seien in ähnlicher Weise $\mathfrak{y}_j$ ($j = 1, 2, \ldots, n_y$) die n_y Ausgänge für die Outputs y_j, deren jeder eine endliche Anzahl, $v_j > 0$, verschiedener Werte annehmen kann. Die Anzahl unterscheidbarer Outputzustände ist dann:

$$Y = \prod_{j=1}^{n_y} v_j. \tag{5}$$

Ein bestimmter Outputzustand $y(t)$ zur Zeit t (oder kurz y) besteht dann in der Identifizierung der Werte y_i in allen n_y Ausgängen $\mathfrak{y}_i$ in diesem „Moment":

$$y(t) \equiv y = \{y_i\}. \tag{6}$$

Sei schließlich Z die Anzahl interner Zustände z, die zum Zwecke dieser Darlegung (wenn nicht anders angegeben) als nicht weiter analysierbar angesehen werden. Die Werte von z können folglich einfach als die natürlichen Zahlen von 1 bis Z angenommen werden, und ein besonderer Outputzustand $z(t)$ zum Zeitpunkt t (oder kurz z) besteht in der Identifizierung des Wertes von z in diesem „Moment":

$$z(t) \equiv z. \tag{7}$$

Jeder dieser „Momente" soll ein endliches Zeitintervall, Δ, dauern, in dem die Werte aller Variablen x, y, z identifiziert werden können. Danach, d. h. zur Zeit $t + \Delta$, nehmen sie Werte $x(t + \Delta)$, $y(t + \Delta)$, $z(t + \Delta)$ (oder kurz x', y', z') an; ihre Werte in der vorausgegangenen Periode $t - \Delta$ waren $x(t - \Delta)$, $y(t - \Delta)$, $z(t - \Delta)$ (oder kurz x^*, y^*, z^*).

Nachdem wir nun die Variablen definiert haben, die in der Maschine wirksam sind, können wir die Operationen über diesen Variablen definieren. Diese sind von zweierlei Art und können auf verschiedene Weisen bestimmt werden. Die gängigste Verfahrensweise besteht darin, zuerst eine „Antriebsfunktion" zu definieren, die

Fortsetzung Fußnote 1

einem Experimentator identifiziert werden kann, der wünscht, daß diese Konfiguration von dem Organismus als ein „Stimulus" verarbeitet wird, kann nicht einfach *als solche* als ein „Inputzustand" der Maschine angenommen werden! Ein derartiger Stimulus kann für den Experimentator stimulierend sein, vom Organismus aber überhaupt nicht beachtet werden. Andererseits kann ein Inputzustand von der Maschine nicht ignoriert werden, es sei denn diese wäre explizit angewiesen worden, eben dies zu tun. Genauer gesprochen, muß die Verteilung der Aktivität der afferenten Nervenfasern als Input angenommen werden, und in ähnlicher Weise die Verteilung der Aktivität der efferenten Fasern als Output des Systems.

in jedem Augenblick den Outputzustand festlegt, und zwar auf der Basis eines gegebenen Inputzustandes und des in diesem Augenblick gegebenen internen Zustandes:

$$y = f_y(x, z). \tag{8}$$

Obwohl nun diese Antriebsfunktion f_y bekannt sein und der Zeitverlauf der Inputzustände x durch den Experimentator gesteuert werden mag, bleiben die Outputzustände y im Ablauf der Zeit so lange unvorhersehbar, als die Werte von z, d.h. die internen Zustände der Maschine, noch nicht bestimmt sind. Es gibt eine große Vielfalt von Möglichkeiten für die Bestimmung des Zeitverlaufs von z in Abhängigkeit von x, von y oder von anderen neu zu definierenden internen oder externen Variablen. Die nützlichste Bestimmung für unsere Zwecke ist, z rekursiv als von vorausgegangenen Tatbeständen abhängige Variable zu definieren. Wir definieren daher die „Zustandsfunktion" f_z der Maschine wie folgt:

$$z = f_z(x^*, z^*), \tag{9a}$$

oder auf andere und äquivalente Weise:

$$z' = f_z(x, z). \tag{9b}$$

Dies bedeutet: der gegenwärtige interne Zustand der Maschine ist eine Funktion ihres vorausgegangenen internen Zustandes sowie ihres vorausgegangenen Inputzustandes; oder in anderer und äquivalenter Weise: der nächste interne Maschinenzustand ist eine Funktion sowohl ihrer gegenwärtigen internen Zustände als auch ihrer Inputzustände.

Das Verhalten der Maschine, d.h. ihre Outputsequenz, ist durch die drei Mengen von Zuständen, {x}, {y}, {z}, und die zwei Funktionen f_y und f_z vollständig determiniert, wenn die Inputsequenz gegeben ist.

Eine derartige Maschine wird als sequentielle, zustandsdeterminierte, „nicht-triviale" Maschine bezeichnet, und in Bild 3a sind die Relationen zwischen ihren verschiedenen Teilen schematisch dargestellt.

Eine solche nicht-triviale Maschine reduziert sich auf eine „triviale" Maschine, wenn sie auf Veränderungen der internen Zustände nicht reagiert, oder wenn die internen Zustände sich nicht ändern (Bild 3b):

$$z' = z = z_0 = \text{konstant} \tag{10a}$$

$$y = f_y(x, \text{konstant}) = f(x). \tag{10b}$$

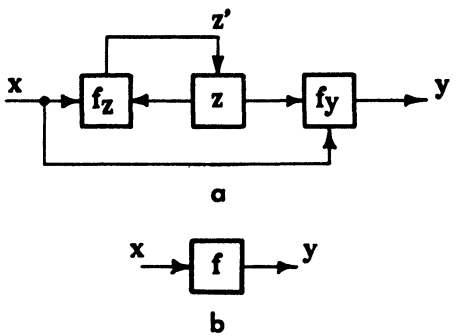


Bild 3
Signalfluß in einer Maschine mit endlich vielen Zuständen (a); Input-Output-Relation in einer trivialen Maschine (b).

Mit anderen Worten, eine triviale Maschine koppelt in deterministischer Weise einen bestimmten Inputzustand mit einem bestimmten Outputzustand, oder, in der Sprache der naiven Reflexologen, einen bestimmten Stimulus mit einer bestimmten Reaktion.

Da der Begriff des „internen Zustands" für die Differenzierung zwischen trivialen und nicht-trivialen Maschinen entscheidend ist, möchte ich nun verschiedene formale Interpretationen dieser Zustände geben, um sie aus der Grauzone der „Unanalysierbarkeit" herauszuholen.

Zunächst mag es scheinen, daß man diese mysteriösen Zustände mit Hilfe eines Kunstgriffs loswerden kann, indem man die Antriebsfunktion f_y in rekursiver Weise definiert. Wie wir jedoch sofort sehen werden, treten diese Zustände dann nur in anderer Form auf.

Nehmen wir einmal die Antriebsfunktion [Gleichung (8)] zur Zeit t, und einen Schritt später (t + Δ):

$$\begin{aligned} y &= f_y(x, z) \\ y' &= f_y(x', z'), \end{aligned} \tag{8'}$$

und nehmen wir an, es gibt eine „inverse Funktion" zu f_y:

$$z = \phi_y(x, y). \tag{11}$$

Wir fügen nun die Zustandsfunktion [Gleichung (9b)] für z' in Gleichung (8') ein und ersetzen z durch Gleichung (11):

$$y' = f_y(x', f_z(x, \phi_y(x, y)) = F_y^{(1)}(x', x, y), \tag{12}$$

oder in anderer und äquivalenter Weise:

$$y = F_y^{(1)}(x, x^*, y^*). \tag{13}$$

Durch Gleichung (13) ist y* jedoch in rekursiver Weise gegeben:

$$y^* = F_y^{(1)}(x^*, x^{**}, y^{**}), \tag{13*}$$

und wenn wir diesen Ausdruck in die Gleichung (13) einfügen, ergibt sich

$$y = F_y^{(2)}(x, x^*, x^{**}, y^{**}),$$

und für n rekursive Schritte

$$y = F_y^{(n)}(x, x^*, x^{**}, x^{***} \ldots x^{(n)*}, y^{(n)*}). \tag{14}$$

Aus diesem Ausdruck geht hervor, daß der Output einer nicht-trivialen Maschine nicht lediglich eine Funktion ihres gegenwärtigen Input ist, sondern vielmehr von der besonderen Sequenz der Inputs, die in die ferne Vergangenheit zurückreicht, sowie von einem Outputzustand in eben dieser fernen Vergangenheit abhängig sein kann. Dies trifft aber nur in einem bestimmten Ausmaß zu – die „Ferne" erstreckt sich nur über Z rekursive Schritte, und außerdem determiniert die Gleichung (14) die Eigenschaften der Maschine nicht in eindeutiger Weise –, und die Abhängigkeit des Verhaltens der Maschine von ihrer Vergangenheit sollte uns daher nicht dazu verführen, in dieses System die Fähigkeit eines Gedächtnisses

zu projizieren, denn es kann im besten Falle nur seinen gegenwärtigen internen Zustand betrachten, der zwar durchaus ein *Ausdruck* der Vergangenheit sein mag, aber dem System keinesfalls die Möglichkeit eröffnet, alles Vergangene zurückzuholen.

Dies läßt sich am leichtesten einsehen, wenn man Gleichung (13) in ihrer vollständigen rekursiven Form für eine lineare Maschine schreibt (mit x und y nun als reellen Zahlen):

$$y(t + \Delta) - ay(t) = bx(t), \tag{15a}$$

oder in der analogen Differentialform mit der Erweiterung $y(t + \Delta) = y(t) + \Delta dy/dt$:

$$\frac{dy}{dt} - \alpha y = x(t), \tag{15b}$$

mit den entsprechenden Lösungen

$$y(n\Delta) = a^n \left[y(0) + b \sum_{i=0}^{n} a^{-i} x(i\Delta) \right], \tag{16a}$$

und

$$y(t) = e^{\alpha t} \left[y(0) + \int_0^t e^{-\alpha \tau} x(\tau)\, d\tau \right]. \tag{16b}$$

Diese Ausdrücke machen klar, daß der Verlauf der Ereignisse, der durch $x(i\Delta)$ (oder $x(\tau)$) repräsentiert ist, „hinausintegriert" wird und nur in einem additiven Term manifest bleibt, der sich nichtsdestoweniger im Laufe der Zeit verändert.

Die Tatsache, daß sich diese simple Maschine als unbrauchbar erweist, das Gedächtnis zu erklären, sollte uns nicht darin hindern, sie als ein möglicherweise nützliches Element eines Systems zu betrachten, das sich tatsächlich erinnern kann.

In den gegebenen Beispielen haben die internen Zustände z es der Maschine ermöglicht, ihre Vergangenheit – wie geringfügig auch immer – zu berücksichtigen; wir werden nun eine Interpretation der internen Zustände z geben, indem wir sie als Selektor einer spezifischen Funktion aus einer Menge mehrwertiger logischer Funktionen auffassen. Dies läßt sich am leichtesten einsehen, wenn die Antriebsfunktion f_y in Form einer Tabelle geschrieben wird.

Seien a, b, c ... X die Inputwerte x; α, β, γ ... Y die Outputwerte y; und 1, 2, 3 ... Z die Werte der internen Zustände. Eine bestimmte Antriebsfunktion f_y ist dann definiert, wenn allen Paaren {xz} ein angemessener Wert von y zugeordnet wird. Dies wird durch Tabelle 1 dargestellt.

Tabelle I Berechnung von Z logischen Funktionen $F_z(x)$ über die Inputs x

z	1	1	1	... 1	2	2	2	... 2	...	Z	Z	Z	... Z
x	a	b	c	... X	a	b	c	... X	...	a	b	c	... X
y	γ	α	β	... δ	α	γ	β	... ϵ	...	β	ϵ	γ	... δ
									...				

Klarerweise wird bei z = 1 eine bestimmte logische Funktion $y = F_1(x)$ definiert, die y mit x verknüpft; bei z = 2 wird eine andere logische Funktion, $y = F_2(x)$ definiert, und allgemein wird für jedes z eine logische Funktion $y = F_z(x)$ definiert.

Die Antriebsfunktion f_y kann daher in folgender Weise neu geschrieben werden:

$$y = F_z(x), \tag{17}$$

und das bedeutet, daß diese Maschine eine neue logische Funktion F_z über ihre Inputs x berechnet, wenn ihre internen Zustände z sich entsprechend der Zustandsfunktion $z' = f_z(x, z)$ verändern.

Oder mit anderen Worten, immer dann, wenn sich z verändert, wird die Maschine zu einer *anderen* trivialen Maschine.

Diese Beobachtung kann wichtig sein, den fundamentalen Unterschied zwischen nicht-trivialen und trivialen Maschinen zu erfassen und die Bedeutung dieses Unterschieds für eine Theorie des Verhaltens zu würdigen, sie erlaubt uns aber auch, die Anzahl der internen Zustände zu berechnen, die die Veränderung des *modus operandi* dieser Maschine bewirken kann.

Gehen wir nun nach dem Modell vor, die Anzahl $\mathfrak{N}$ der logischen Funktionen als die Anzahl der Zustände der abhängigen Variablen erhoben zur Potenz der Anzahl der Zustände der unabhängigen Variablen zu berechnen:

$$\mathfrak{N} = \boxed{\text{(Anzahl der Zustände der abhängigen Variablen)}}^{\boxed{\text{(Anzahl der Zustände der unabhängigen Variablen)}}} \tag{18}$$

dann ergibt sich als die Anzahl möglicher trivialer Maschinen, die y mit x verknüpfen:

$$\mathfrak{N}_T = Y^X. \tag{19}$$

Dies ist jedoch die größte Anzahl interner Zustände, die tatsächlich eine Veränderung in der Funktion $F_z(x)$ bewirken kann, denn jeder zusätzliche Zustand muß mit einer Funktion gepaart werden, der bereits ein Zustand zugeordnet worden ist, so daß derartige zusätzliche interne Zustände redundant oder zumindest ununterscheidbar sind. Es gilt daher

$$Z \leqslant Y^X.$$

Die Gesamtzahl der Antriebsfunktionen $f_y(x, z)$ ist

$$\mathfrak{N}_D = Y^{XZ}, \tag{20}$$

ihr größter Wert ist daher

$$\mathfrak{N}_D = Y^{XY^X}. \tag{21}$$

In ähnlicher Weise ergibt sich die Anzahl der Zustandsfunktionen $f_z(z, x)$

$$\mathfrak{N}_S = Z^{X \cdot Z}, \tag{22}$$

und ihr größter effektiver Wert ist

$$\mathfrak{N}_S = Y^{X \cdot XY^X} = [\mathfrak{N}_D]^X. \tag{23}$$

Diese Zahlen wachsen sehr schnell zu meta-astronomischen Größen auch für Maschinen an, die nur höchst bescheidende Ansprüche erheben.

Wir wollen annehmen, daß eine Maschine nur einen zweiwertigen Output hat ($n_y = 1$; $v_y = 2$; $y = \{0; 1\}$; $Y = 2$), und daß sie n zweiwertige Inputs hat ($n_z = n$; $v_x = 2$; $x = \{0; 1\}$; $X = 2^n$). Tabelle 2 stellt die Anzahl der effektiven internen Zustände, die Anzahl der möglichen Antriebsfunktionen, und die Anzahl effektiver Zustandsfunktionen für Maschinen dar, die bis zu vier „Afferenten" entsprechend den folgenden Gleichungen aufweisen:

$$Z = 2^{2^n}$$
$$\mathfrak{N}_D = 2^{2^{n+n}}$$
$$\mathfrak{N}_S = 2^{2^{2n+2n}}$$

Tabelle 2 Die Anzahl der effektiven Zustände Z, die Anzahl der möglichen Antriebsfunktionen $_D$ und die Anzahl der effektiven Zustandsfunktionen $_S$ für Maschinen mit einem zweiwertigen Output und mit bis zu vier zweiwertigen Inputs

n	Z	$\mathfrak{N}_D$	$\mathfrak{N}_S$
1	4	256	65536
2	16	$2 \cdot 10^{19}$	$6 \cdot 10^{76}$
3	256	10^{600}	$300 \cdot 10^{4 \cdot 103}$
4	65536	$300 \cdot 10^{4 \cdot 103}$	$1600 \cdot 10^{7 \cdot 104}$

Diese rasch ansteigenden Zahlen drücken aus, daß schon auf der molekularen Ebene ohne viel Aufwand eine rechnerische Mannigfaltigkeit am Werke sein kann, die alle unsere Vorstellungen übersteigt. Ganz offensichtlich beweist die große Vielfalt der Ergebnisse genetischer Errechnungen, wie sie sich in der Mannigfaltigkeit der Lebensformen auch nur innerhalb einer einzigen Art manifestiert, derartige Möglichkeiten. Eine genauere Erörterung dieser Möglichkeiten soll jedoch dem nächsten Abschnitt vorbehalten bleiben.

2 *Interagierende Maschinen*

Wir wollen nun den allgemeineren Fall erörtern, in dem zwei oder mehrere solche Maschinen miteinander interagieren. Wenn bestimmte Aspekte des Verhaltens eines Organismus durch eine Maschine mit endlich vielen Zuständen modelliert werden können, dann könnte die Interaktion des Organismus mit seiner Umwelt ein solcher Fall sein, wenn die Umwelt in gleicher Weise durch eine Maschine mit endlich vielen Zuständen repräsentiert werden kann. In der Tat bilden derartige Interaktionen zwischen zwei Maschinen das gängige Modell des Verhaltens von Lebewesen in experimentellen Lernsituationen, wobei die Komplexität der Situation gewöhnlich dadurch verringert wird, daß als experimentelle Umwelt eine triviale Maschine eingesetzt wird. Das „Kriterium" dieser Lernexperimente wird vom Lebewesen angeblich dann erfüllt, wenn es dem Experimentator gelungen ist, es aus einer nicht-trivialen Maschine in eine triviale Maschine zu transfor-

mieren, so daß aufgrund der Experimente nur noch zwei triviale Maschinen miteinander interagieren.

Wir wollen die zur Umwelt (E) gehörenden Quantitäten mit römischen Lettern bezeichnen, die zum Organismus (Ω) gehörenden mit den entsprechenden griechischen Buchstaben. So lange E und Ω unabhängig sind, legen sechs Gleichungen ihr Schicksal fest. Die vier „Maschinengleichungen", zwei für jedes System,

$$E: \quad y = f_y(x, z) \tag{24a}$$

$$z' = f_z(x, z) \tag{24b}$$

$$\Omega: \quad \eta = f_\eta(\xi, \zeta) \tag{25a}$$

$$\zeta' = f_\zeta(\xi, \zeta) \tag{25b}$$

und die zwei Gleichungen, die den Ereignisverlauf an den „Rezeptoren" der beiden Systeme beschreiben:

$$x = x(t); \quad \xi = \xi(t). \tag{26a, b}$$

Wir lassen nun diese beiden Systeme miteinander interagieren, indem wir den (um einen Schritt verzögerten) Output jeder Maschine mit dem Input der anderen verknüpfen. Die Verzögerung soll eine Art „Reaktionszeit" (Berechnungszeit) jedes Systems auf einen gegebenen Input (Stimulus, Ursache) (vgl. Bild 4) repräsentieren. Mit diesen Verbindungen sind nun die folgenden Relationen zwischen den externen Variablen der beiden Systeme hergestellt:

$$x' = \eta = u'; \quad \xi' = y = v', \tag{27a, b}$$

wobei die neuen Variablen u und v die von $\Omega \rightarrow E$ bzw. $E \rightarrow \Omega$ übertragenen „Botschaften" darstellen. Wenn wir in den Gleichungen (24) und (25) entsprechend Gleichung (27) x, y, η, ξ, durch u und v ersetzen, ergibt sich:

$$\begin{aligned} v' &= f_y(u, z); \quad & u' &= f_\eta(v, \zeta) \\ z' &= f_z(u, z); \quad & \zeta' &= f_\zeta(v, \zeta). \end{aligned} \tag{28}$$

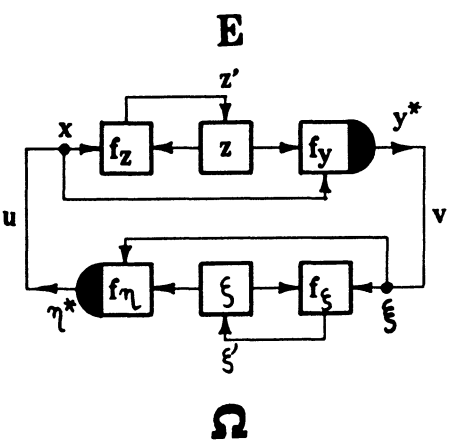


Bild 4
Zwei Maschinen mit endlich vielen Zuständen (*E*) (Ω), die über Verzögerungen (schwarze Halbkreise) verbunden sind.

Dies sind vier rekursive Gleichungen für die vier Variablen u, v, z, ζ, und wenn die vier Funktionen f_y, f_z, f_η, f_ζ gegeben sind, ist das Problem, u(t), v(t), z(t), ζ(t) zu „lösen", d.h. diese Variablen explizit als Zeitfunktionen auszudrücken, ein rein mathematisches. Mit anderen Worten, das „Meta-System" (EΩ), gebildet aus den Subsystemen E und Ω, ist sowohl physikalisch als auch mathematisch „geschlossen", und sein Verhalten ist für alle Zeiten vollständig determiniert. Wenn außerdem in einem gegebenen Moment, etwa t = 0 (Ausgangsbedingungen), die Werte aller Variablen u(0), v(0), z(0), ζ(0) bekannt sind, ist das System auch vollständig vorhersagbar. Da dieses Meta-System keinen Input hat, läuft es nach diesen seinen Regeln so lange einfach weiter, bis es schließlich irgendwann ein statisches oder dynamisches Gleichgewicht erreicht, je nachdem wie seine Regeln und die Ausgangsbedingungen aussehen.

Der allgemeine Fall des Verhaltens solcher Systeme ist mit Hilfe von Computersimulationen eingehend untersucht worden (Walker 1965; Ashby/Walker 1966; Fitzhugh 1963), während die Lösungen für die Gleichungen (28) im linearen Fall auf direktem Wege erzielt werden können, besonders dann, wenn angenommen werden darf, daß sich die Rekursionen über infinitesimal kleine Schritte erstrecken:

$$w' = w(t + \Delta) = w(t) + \Delta \frac{dw}{dt}. \tag{29}$$

Unter diesen Bedingungen werden die vier Gleichungen (28) zu

$$w_i = \sum_{j=1}^{4} a_{ij} w_j, \tag{30}$$

wobei die w_i (i = 1, 2, 3, 4) nunmehr die reellen Zahlen sind und die vier entsprechenden Variablen ersetzen, w die erste abgeleitete Größe mit Bezug auf die Zeit repräsentiert, und die 16 Koeffizienten a_{ij} (i, j = 1, 2, 3, 4) die vier entsprechenden linearen Funktionen definieren. Dieses System simultaner linearer Differentialgleichungen erster Ordnung wird gelöst durch

$$w_i(t) = \sum_{j=1}^{4} A_{ij} e^{\lambda j t} \tag{31}$$

wobei λ_j die Wurzeln der Determinante sind

$$|a_{ij} - \delta_{ij}\lambda| = 0$$

$$\delta_{ij} = \begin{cases} 1 \ldots i = j \\ 0 \ldots i \neq j \end{cases} \tag{32}$$

und die A_{ij} von den Ausgangsbedingungen abhängen. Je nachdem, ob die λ_j komplexe, negative oder positive reelle Zahlen sind, wird das System letztlich oszillieren, absterben, oder explodieren.[2])

---

2) Dieses Ergebnis ist natürlich unmöglich in einer Maschine mit endlich vielen Zuständen. Es ergibt sich hier nur, weil die diskreten und finiten Variablen u, v, z, ζ durch w_i ersetzt worden sind, die kontinuierliche und unbegrenzte Quantitäten darstellen.

Eine eingehende Erörterung der verschiedenen Verhaltensweisen solcher Systeme würde den Rahmen dieser Übersicht sprengen. Es sollte allerdings festgehalten werden, daß in allen Fällen als Verhaltensmerkmal eine anfängliche Übergangsphase auftritt, die eine sehr große Zahl von Zuständen durchlaufen kann, bis schließlich ein Zustand erreicht wird, der eine stabile zyklische Bahn festlegt: das dynamische Gleichgewicht. Form und Länge sowohl der transitorischen als auch der endgültigen Gleichgewichtsphasen sind von den Ausgangsbedingungen abhängig, eine Tatsache, die Ashby (1956) solche Systeme als „multistabil" bezeichnen ließ. Da gewöhnlich eine große Menge von Ausgangsbedingungen auf einen einzigen Gleichgewichtszustand abgebildet wird, läßt sich dieser Gleichgewichtszustand als eine *dynamische Repräsentation* einer Menge von Ereignissen auffassen, und in einem multistabilen System kann daher jeder Zyklus als eine „Abstraktion" dieser Ereignisse gesehen werden.

Mit Hilfe dieser Konzepte wollen wir nun sehen, was sich aus einem typischen Lernexperiment (z.B. John u.a. 1969) ableiten läßt, in dem das Versuchstier in einem Y-Labyrinth die Wahl ($\xi_0 \equiv$ C für „Wahl") zwischen zwei Handlungen ($\eta_1 \equiv$ L, „links abbiegen"; $\eta_2 \equiv$ R, „rechts abbiegen") angeboten bekommt. Auf diese Handlungen hin reagiert die Umwelt E, eine triviale Maschine, mit neuen Inputs für das Tier ($\eta_1 = x_1' \rightarrow y_1' = \xi_1'' \equiv$ S, „Schock"; oder $\eta_2' = x_2' \rightarrow y_2' = \xi_2'' \equiv$ F, „Nahrung"), die in diesem Schmerz ($\eta_3 \equiv$ „–") oder Lust ($\eta_4 \equiv$ „+") erzeugen. Diese Reaktionen veranlassen E, das Tier in die ursprüngliche Wahlsituation ($\xi_0 \equiv$ C) zurückzuversetzen.

Betrachten wir nun die einfache Überlebensstrategie, die in das Tier eingebaut ist, und durch die es unter neutralen und lustvollen Bedingungen seine internen Zustände [$\zeta' = \zeta$, für (Cζ) und (Fζ)] aufrechterhält, während es sie unter Schmerzbedingungen ändert [$\zeta' \neq \zeta$, für (Sζ)]. Wir wollen acht innere Zustände annehmen ($\zeta = i$, $i = 1, 2, 3, \ldots, 8$).

Durch diese Regeln ist das gesamte System (ΩE) bestimmt und sein Verhalten vollständig determiniert. Der Einfachheit halber sind die drei Funktionen $f_y = f$ für die triviale Maschine E, $f\eta$ und f_ζ für Ω, in den Tabellen 3a, b, c aufgelistet.

Tabelle 3a

$y = f(x)$

x (= η^*)	y (= ξ')
L	S
R	F
–	C
+	C

Tabelle 3b

$\eta = f_\eta(\xi, \zeta)$

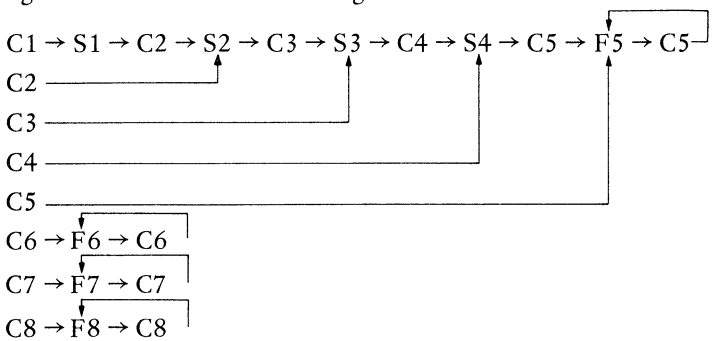
η (= x′)		ζ 1	2	3	4	5	6	7	8
ξ (= y*)	C	L	L	L	L	R	R	R	R
	S	–	–	–	–	–	–	–	–
	F	+	+	+	+	+	+	+	+

Tabelle 3c

$\zeta' = f_\zeta(\xi, \zeta)$

ζ'		ζ 1	2	3	4	5	6	7	8
ξ (= y*)	C	1	2	3	4	5	6	7	8
	S	2	3	4	5	6	7	8	1
	F	1	2	3	4	5	6	7	8

Mit Hilfe dieser Tabellen lassen sich die acht Verhaltensbahnen des (ΩE)-Systems entsprechend den acht Ausgangsbedingungen beschreiben. Dies ist im folgenden dargestellt, wobei lediglich die Werte der Paare ξζ angegeben sind, wie sie aufgrund der Reaktionen des Organismus bzw. der Umwelt aufeinander folgen.



Diese Bahnen zeigen das Verhalten genau, wie es vorhin gekennzeichnet wurde: die Anfangsübergänge hängen in ihrer Länge von den Ausgangsbedingungen ab; letztlich ergibt sich dynamisches Gleichgewicht, das zwischen den beiden externen Zuständen ohne Veränderung des internen Zustandes wechselt. Das ganze System und seine Teile sind zu trivialen Maschinen geworden. Da man auch mit maximaler semantischer Toleranz nicht behaupten kann, daß eine triviale Maschine ein Gedächtnis hat, muß man sich fragen, was gemessen werden soll, wenn sie in dieser Phase aufgemacht wird, und wenn ihre internen Mechanismen untersucht werden. Möchte man ihre gegenwärtige Funktionsweise prüfen? Oder möchte man feststellen, wie sehr sich die Maschine seit früheren Prüfungen verändert hat? Im besten Falle handelt es sich dabei um Tests des Gedächtnisses des Experimentators, ob die Maschine aber irgendwelche Veränderungen selbst verarbeiten kann, kann prinzipiell nicht durch Experimente erfahren werden, deren Konzeption eben die Qualität eliminiert, die sie messen sollen.

3 *Probabilistische Maschinen*

Dieses Dilemma läßt sich noch deutlicher machen, wenn wir für einen Augenblick die Perspektive der statistischen Lerntheorie einnehmen (Skinner 1959; Estes 1959; Logan 1959). Hier wird entweder die Vorstellung interner Zustände verworfen oder die Existenz interner Zustände ignoriert. Immer dann aber, wenn die Gesetze ignoriert werden, die Ursachen mit Wirkungen verknüpfen, sei es aus Ignoranz oder durch bewußte Entscheidung, wird die Theorie zu einer Theorie von Wahrscheinlichkeiten.

Wenn wir im vorausgegangenen Beispiel den Ausgangszustand nicht kennen, dann stehen die Chancen 50 zu 50, daß sich das Tier beim ersten Versuch entweder nach links oder nach rechts wendet. Nach einem Durchgang ist die Wahrscheinlichkeit, daß es sich nach rechts wendet, 5/8 usw., bis das Tier sich aus einer „probabilistischen (nicht-trivialen) Maschine" in eine „deterministische (triviale) Maschine" verwandelt hat, und sich von da an stets nach rechts wendet. Während nun der Verfechter einer statistischen Lerntheorie die sich verändernden Wahrscheinlichkeiten in jedem der nachfolgenden Versuche zu einem „ersten Prinzip" erheben wird, ist dies für den Betreiber einer Maschine mit endlich vielen Zustän-

den eine ganz offensichtliche Konsequenz der Wirkung gewisser Inputs auf die internen Zustände seiner Maschine: sie werden durch die Koppelung mit „schmerzhaften Inputs" unzugänglich. Der ganze mathematische Apparat der statistischen Lerntheorie kann faktisch auf das Paradigma reduziert werden, unter Beachtung gewisser Nicht-Ersetzungsregeln Kugeln von verschiedener Farbe aus einem Gefäß zu ziehen.

Gegeben sei ein Gefäß mit Kugeln von m verschiedenen Farben, die mit 0, 1, 2, ..., (m − 1) etikettiert sind. Gewisse noch nicht definierte Regeln erlauben oder verbieten es, eine bestimmte farbige Kugel in das Gefäß zurückzugeben, sobald sie herausgezogen worden ist. Betrachten wir nun die Ergebnisse einer Sequenz von n Ziehungen, einer „n-Sequenz", als eine n-stellige m-adische Zahl (z.B. m = 10; n = 12).

v = 1 5 7 3 0 2 1 8 6 2 1 4
↑ ↑
zuletzt zuerst
gezogen gezogen.

Damit ist klar, daß es

$$\mathfrak{N}(n, m) = m^n$$

verschiedene n-Sequenzen gibt. Eine *bestimmte* n-Sequenz soll eine ν-Zahl heißen, d.h.:

$$0 \leqslant \nu(m, n) = \sum_{i=1}^{n} j(i)\, m^{(i-1)} \leqslant m_{-1}^{n}, \tag{33}$$

wobei $0 \leqslant j(i) \leqslant (m = 1)$ das mit j etikettierte Ergebnis des i-ten Versuchs darstellt.

Die Wahrscheinlichkeit einer *bestimmten* n-Sequenz (dargestellt durch eine ν-Zahl) ist somit

$$p_n(\nu) = \prod_{i=1}^{n} p_i[j(i)], \tag{34}$$

wobei $p_i[j(i)]$ die Wahrscheinlichkeit dafür angibt, daß die mit j etikettierte Farbe beim i-ten Versuch in Übereinstimmung mit der spezifischen ν-Zahl, wie sie durch Gleichung (33) definiert wird, auftritt.

Da nach jedem Versuch mit einem „Nicht-zurückgeben"-Ergebnis alle Wahrscheinlichkeiten geändert werden, wird die Wahrscheinlichkeit eines Ereignisses beim n-ten Versuch von dem „Verlaufsweg", d.h. von der Vergangenheit bzw. der Geschichte der Ereignisse abhängen, die zu diesem Ereignis geführt haben. Da es m^{n-1} mögliche Verlaufswege gibt, die dem Ziehen von j beim n-ten Versuch vorausgehen können, ergibt sich die Wahrscheinlichkeit dieses Ereignisses wie folgt:

$$p_n(j) = \sum_{\nu=0}^{m^{n-2}-1} p_n(j \cdot m^{n-1} + \nu(n-1, m)),$$

wobei $j \cdot m^{n-1} + \nu(n-1, m)$ eine $\nu(n, m)$-Zahl repräsentiert, die mit j beginnt.

Davon läßt sich eine nützliche Rekursion ableiten. Seien j* die Farben der Kugeln, die *nicht* ersetzt werden, wenn sie gezogen sind, und seien j die anderen Kugeln. Seien weiterhin n_{j*} und n_j die Anzahl der vorausgegangenen Versuche, bei denen j* bzw. j gezogen wurden ($\Sigma n_{j*} + \Sigma n_j = n - 1$), dann ist die Wahrscheinlichkeit, j (oder j*) beim n-ten Versuch mit einem Verlaufsweg von Σn_{j*} Entnahmen zu ziehen

$$p_n(j) = \frac{N_j}{N - \Sigma n_{j*}} \cdot p_{n-1}(\Sigma n_{j*}) \tag{35a}$$

und

$$p_n(j^*) = \frac{N_{j*} - n_{j*}}{N - \Sigma n_{j*}} \cdot p_{n-1}(\Sigma n_{j*}), \tag{35b}$$

wobei $N = \Sigma N_j + \Sigma N_{j*}$ die anfängliche Anzahl an Kugeln ist, und N_j und N_{j*} die anfängliche Anzahl an Kugeln mit den Farben j bzw. j*.

Wir wollen mit N Kugeln in einem Gefäß beginnen, wobei N_w davon weiß sind und $(N - N_w)$ schwarz. Wird eine weiße Kugel gezogen, dann wird sie wieder in das Gefäß zurückgegeben; eine schwarze Kugel wird jedoch beiseite gelegt. Mit „weiß" ≡ 0 und „schwarz" ≡ 1 wäre eine bestimmte n-Sequenz (n = 3):

$$\nu(3, 2) = 1\ 0\ 1,$$

und ihre Wahrscheinlichkeit:

$$p_3(1\ 0\ 1) = \frac{N - N_w - 1}{N - 1} \cdot \frac{N_w - 1}{N - 1} \cdot \frac{N - N_w}{N}.$$

Die Wahrscheinlichkeit, beim dritten Versuch eine schwarze Kugel zu ziehen, ist somit:

$$p_3(1) = p_3(1\ 0\ 0) + p_3(1\ 0\ 1) + p_3(1\ 1\ 0) + p_3(1\ 1\ 1).$$

Wir wollen nun die Wahrscheinlichkeit feststellen, beim n-ten Versuch eine weiße Kugel zu ziehen. Diese Wahrscheinlichkeit sei durch p(n) bezeichnet, die Wahrscheinlichkeit, eine schwarze Kugel zu ziehen mit $q(n) = 1 - p(n)$.

Versucht man [mit Gleichung (35)], durch iterative Approximation zu erreichen, daß Versuchsfolgen von der Länge m verlaufsunabhängig werden [$p_i(j) = p_1(j)$], dann erhält man eine Approximation erster Ordnung an eine Rekursion in p(n):

$$p(n) = p(n - m) + \frac{m}{N} q(n - m) \tag{36}$$

oder für $m = n - 1$ (gültig für $p(1) \approx 1$, und $n/N \ll 1$):

$$p(n) = p(1) + \frac{n - 1}{N} q(1) \tag{37}$$

und für m = 1 (gültig für $p(1) \approx 1$):

$$p(n) = p(n - 1) + \frac{1}{N} q(n - 1). \tag{38}$$

Eine zweite Approximation verändert den obigen Ausdruck zu

$$p(n) = p(n-1) + \theta q(n-1) \tag{39}$$

wobei $\theta = \theta(N, N_w)$ in allen Versuchen konstant ist.

Damit ergibt sich

$$p(n) - p(n-1) = \Delta p = \theta(1-p) \tag{40}$$

und mit dem Limes

$$\lim_{\Delta n \to 0} \frac{\Delta p}{\Delta n} = \frac{dp}{dn}$$

ergibt sich

$$\frac{dp}{dn} = \theta(1 - p(n))$$

mit der Lösung

$$p(n) = 1 - (1 - p_0)\, e^{-\theta n}. \tag{41}$$

Dies wiederum ist eine Approximation für $p \approx 1$ von

$$p(n) = \frac{p_0}{p_0 + (1 - p_0)\, e^{-\theta n}}, \tag{42}$$

was die Lösung darstellt für

$$\frac{dp}{dn} = \theta p(1-p) \tag{43}$$

oder, rekursiv ausgedrückt, für

$$p(n) = p(n-1) + \theta p(n-1) \cdot q(n-1). \tag{44}$$

Bild 5a vergleicht die Wahrscheinlichkeiten p(n), beim n-ten Versuch eine weiße Kugel zu ziehen, wie sie durch Approximation [Gleichung (42)] (durchgezogene Kurven) berechnet wurden, mit den exakten Werten, die auf einem IBM 360/50-System mit Hilfe eines Programms berechnet wurden, das Herr Atwood freundlicherweise zur Verfügung gestellt hat, und zwar für ein Gefäß, das von Anfang an vier Kugeln (N = 4) enthielt, bzw. für die drei Fälle, in denen eine, zwei oder drei Kugeln weiß sind ($N_w = 1; N_w = 2; N_w = 3$). Die Entropie[3] H(n) in Bits pro Versuch, wie sie diesen Fällen entspricht, wird in Bild 5b gezeigt, und man kann feststellen, daß sie zwar für gewisse Fälle [$p(1) \leqslant 0{,}5$] im Verlauf des Spiels ein Maximum erreicht, daß sie aber in all den Fällen verschwindet, in denen Gewißheit des Ergebnisses approximiert wird [$p(n) \to 1$].

Auch wenn sich diese Skizze der Wahrscheinlichkeiten ausschließlich mit Gefäßen, Kugeln und Ziehungen beschäftigt hat, werden Verfechter der statistischen Lerntheorie in den Gleichungen (39), (41) und (42) die Basisaxiome dieser

[3] Oder: „Betrag der Unsicherheit", „Informationsbetrag", der durch das Ergebnis jedes Versuchs erhalten wird, definiert durch $-H(n) = p(n) \log_2 p(n) + q(n) \log_2 q(n)$.

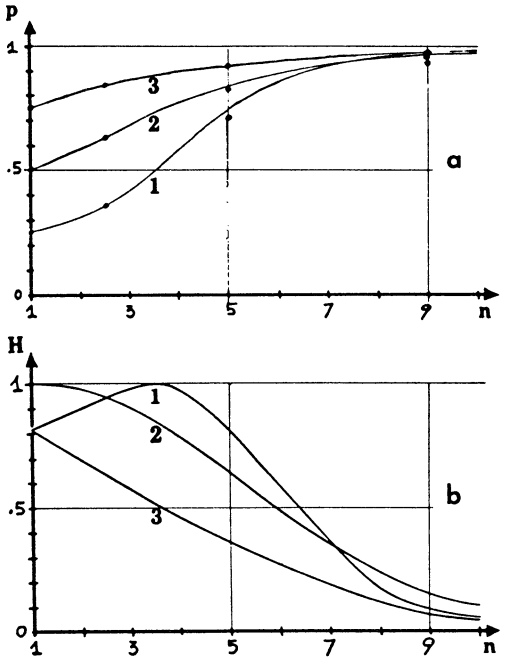


Bild 5
Wahrscheinlichkeit, beim *n*-ten Versuch eine weiße Kugel aus einem Gefäß zu ziehen, in dem sich von Anfang an vier Kugeln befinden, von denen eine, zwei oder drei weiß, die übrigen schwarz sind. Die weißen Kugeln werden stets in den Behälter zurückgegeben, die schwarzen nicht (*a*). Entropie beim *n*-ten Versuch (*b*).

Theorie wiedererkannt haben [Estes 1959; Gleichungen (5), (6) und (9)], und es gibt heute keinen Zweifel, daß Tiere unter den gegebenen experimentellen Bedingungen in der Tat genau die Lernkurven zeichnen, die für diese Bedingungen abgeleitet worden sind.

Da der Formalismus, der für das Verhalten dieser experimentellen Lebewesen gilt, auch für unser Gefäß gilt, stellt sich nun die Frage: Können wir behaupten, daß ein Gefäß lernt? Ist die Antwort „Ja", dann bedarf es ganz offensichtlich für das Lernen keines *Gedächtnisses*, denn unser Gefäß zeigt keinerlei Spuren von schwarzen Kugeln, wenn es schließlich korrekt mit weißen Kugeln auf die „Stimulierung" durch eine Ziehung „reagiert"; ist die Antwort „Nein", dann müssen wir durch Analogie zur Schlußfolgerung kommen, daß in diesen Tierversuchen jedenfalls nichts beobachtet wird, was *Lernen* entspricht.

Um diesem Dilemma zu entgehen, braucht man sich nur daran zu erinnern, daß ein Gefäß eben nichts weiter als ein Gefäß ist, daß es aber Lebewesen sind, die lernen. Lernen findet nun in diesen Experimenten in der Tat auf zwei Ebenen statt. Einmal lernen die experimentellen Lebewesen, sich in einer „gefäßgemäßen" Weise zu verhalten, oder besser, sich auf solche Weise zu verhalten, daß dem Experimentator die Anwendung gefäßgemäßer Kriterien möglich wird. Zum anderen hat der Experimentator etwas über die Tiere erfahren, indem er sie aus nichttrivialen (probabilistischen) Maschinen in triviale (deterministische) Maschinen verwandelt hat. Es ist also die Untersuchung des Experimentators, die uns Hinweise auf Gedächtnis und Lernen liefern muß.

C Maschinen mit endlich vielen Funktionen

1 *Deterministische Maschinen*

Aufgrund unserer Beobachtung richtet sich nun die Frage, wo wir nach Gedächtnis und Lernen suchen sollen, in die genau entgegengesetzte Richtung. Statt in der Umwelt nach Mechanismen zu suchen, die Organismen in triviale Maschinen verwandeln, müssen wir die Mechanismen innerhalb der Organismen feststellen, die diese instandsetzen, ihre Umwelt zu einer trivialen Maschine zu machen.

Nach dieser Formulierung des Problems scheint klar zu sein, daß ein Organismus für die Manipulation seiner Umwelt – irgendwie – eine interne Repräsentation all der Gesetzmäßigkeiten der Umwelt konstruieren muß, die er überhaupt erfassen kann. Die Neurophysiologen wissen seit langem von den abstrahierenden Rechenleistungen, die neuronale Netze von der Rezeptorenebene aufwärts bis zu den höheren Nuklei ausführen (Lettvin u. a. 1959; Maturana u. a. 1968; Eccles u. a. 1967). Mit anderen Worten, es stellt sich hier die Frage, wie anstelle von Zuständen Funktionen berechnet werden, oder wie eine Maschine zu bauen ist, die anstelle numerischer Ergebnisse Programme errechnet. Dies bedeutet, daß wir nach einem Formalismus suchen müssen, der „Maschinen mit endlich vielen *Funktionen*" bewältigt. Ein derartiger Formalismus liegt natürlich eine Ebene höher als der vorhin erörterte, einige seiner wesentlichen Merkmale lassen sich aber dennoch deutlich machen, wenn wir verschiedene der jeweils gültigen Analogien beibehalten.

Unsere Variablen sind nunmehr Funktionen, und da Relationen zwischen Funktionen gewöhnlich als „Funktionale" bezeichnet werden, werde ich nun kurz die wesentlichen Merkmale eines Kalküls rekursiver Funktionale skizzieren.

Nehmen wir ein System wie das in Bild 3a dargestellte und lassen es nunmehr im Unterschied dazu über einer endlichen Menge von Funktionen von zweierlei Art operieren, $\{f_{yi}\}$ und $\{f_{zj}\}$. Diese Funktionen operieren ihrerseits über ihre jeweiligen Mengen von Zuständen $\{y_i\}$ und $\{z_j\}$. Die Operationsregeln für eine derartige Maschine mit endlich vielen Funktionen werden nun exakt entsprechend den Regeln für Maschinen mit endlich vielen Zuständen formuliert. Es ergibt sich:

$$f_y = F_y[x, f_z] \tag{45a}$$

$$f_z' = F_z[x, f_z], \tag{45b}$$

wobei F_y und F_z die Funktionale bedeuten, die aus der gegebenen internen Funktion f_z und einem Input x die Antriebsfunktionen f_y und die sich daraus ergebenden internen Funktionen f_z' generieren. Es ist jedoch zu bemerken, daß der Input hier immer noch ein Zustand ist. Dies deutet auf ein wichtiges Merkmal dieses Formalismus hin, nämlich auf die Einführung eines Bindeglieds zwischen dem Zustandsbereich und dem davon völlig verschiedenen Funktionsbereich. Mit anderen Worten, dieser Formalismus trifft eine Unterscheidung zwischen Entitäten und ihren Repräsentationen und stellt zwischen diesen beiden Bereichen eine Beziehung her.

Mit Hilfe eines Verfahrens, das dem in den Gleichungen (10)–(14) angewandten ähnlich ist, lassen sich die Funktionen vom Typ f_z eliminieren, indem die gegenwärtige Antriebsfunktion als das Ergebnis früherer Zustandsbedingungen abgeleitet wird. Aufgrund gewisser Eigenschaften jedoch, die Funktionale von Funk-

tionen unterscheiden, schließen diese früheren Zustandsbedingungen sowohl Inputzustände als auch Outputfunktionen ein. Es ergibt sich für n rekursive Schritte:

$$f_y = \Phi_y^{(n)}[x, x^*, x^{**}, x^{***}, \ldots, x^{(n)*}; f_y^*, f_y^{**}, \ldots, f_y^{(n)*}]. \tag{46}$$

Vergleicht man diesen Ausdruck mit dem analogen Ausdruck für Maschinen mit endlich vielen Zuständen (14), dann zeigt sich deutlich, daß der Bezug auf vergangene Ereignisse sich hier nicht nur auf jene Ereignisse erstreckt, die die Geschichte der Inputs des Systems ausmachen, $\{x^{(i)*}\}$, sondern auch auf seine Geschichte potentieller Aktionen, $\{f_y^{(i)*}\}$. Wird dieses rekursive Funktional außerdem mit explizitem Zeitbezug gelöst ($t = k\Delta$; $k = 0, 1, 2, 3, \ldots;$) [vgl. Gleichung (16)], dann wird wiederum die Geschichte der Inputs „hinausintegriert"; die Geschichte der potentiellen Aktionen bleibt jedoch intakt, denn es gibt eine Menge von „Eigenfunktionen", die Gleichung (46) genügen. Für $(k\Delta)$ und für die i-te Eigenfunktion gilt explizit:

$$f_y^i(k\Delta) = K_i(k\Delta) \cdot [\pi_i(f_y^{(i)*}) + G_i(x, x^*, x^{**}, \ldots, x^{(n)*})] \qquad i = 1, 2, 3, \ldots, n, \tag{47}$$

wobei K_i und G_i Funktionen von $(k\Delta)$ sind, und wobei der letztere Ausdruck einen Wert ergibt, der von einer Folge von Werten in $x^{(j)*}$ abhängt, die n Schritte lang ist. π_i ist wiederum ein Funktional und repräsentiert die Outputfunktion f_y von i Schritten in der Vergangenheit durch eine andere Funktion.

Auch wenn dieser Formalismus keinen Mechanismus festlegt, der die erforderlichen Rechenoperationen ausführen könnte, liefert er uns zumindest eine adäquate Beschreibung der funktionalen Organisation des Gedächtnisses. Der Zugang zur „vergangenen Erfahrung" ist hier durch die Verfügbarkeit des dem System eigenen *modus operandi* aus früheren Situationen gegeben, und es ist beruhigend, dem Ausdruck (47) entnehmen zu können, daß die subtile Unterscheidung zwischen einer Erfahrung in der Vergangenheit, $(f_y^{(i)*})$, und der gegenwärtigen Erfahrung einer Erfahrung in der Vergangenheit, $[\pi_i(f_y^{(i)*})]$, – d.h. der Unterschied zwischen „Erfahrung" und „Erinnerung" – tatsächlich durch diesen Formalismus angemessen berücksichtigt wird. Das System kann außerdem aufgrund seines Zugangs zu früheren Zuständen seines Funktionierens, und eben nicht aufgrund eines Rückgriffs auf eine gespeicherte Ansammlung von zufälligen Paaren $\{x_i, y_i\}$, die dieses Funktionieren manifestieren, einen Strom von „Daten" errechnen, die mit der vergangenen Erfahrung des Systems konsistent sind. Diese Daten können nun die Outputwerte $\{y_i\}$ dieser Zufallspaare enthalten oder auch nicht. Diesen Preis muß man zahlen, wenn man die Bereiche wechselt, wenn man von Zuständen zu Funktionen und wieder zurück zu Zuständen geht. Dies ist jedoch ein in der Tat sehr geringer Preis angesichts des Gewinns eines unendlich leistungsfähigeren „Speicherungssystems", das die Antwort auf eine Frage *errechnet*, und eben nicht alle Antworten zusammen mit allen möglichen Fragen speichert, um eine Antwort dann bereitzustellen, wenn es die Frage finden kann (von Foerster 1965).

Diese Beispiele mögen genügen, um ohne größere Schwierigkeiten eine weitere Eigenschaft der Maschine mit endlich vielen Funktionen zu interpretieren, die der Maschine mit endlich vielen Zuständen genau analog ist. Eine Maschine mit endlich vielen Funktionen wird ebenso wie eine Maschine mit endlich vielen Zustän-

den bei der Interaktion mit einem anderen System anfänglich Übergangszustände durchlaufen, die von den Ausgangsbedingungen abhängen, und schließlich in dynamisches Gleichgewicht gelangen. Wir haben hier wiederum, wenn es keine interne Veränderung der Funktion gibt ($f'_z = f_z = f_0$), eine „triviale Maschine mit endlich vielen Funktionen" mit ihrer „Zielfunktion" f_0. Es ist leicht zu sehen, daß eine triviale Maschine mit endlich vielen Funktionen einer nicht-trivialen Maschine mit endlich vielen Zuständen äquivalent ist.[4])

Statt nun weitere Eigenschaften der funktionalen Organisation von Maschinen mit einer endlichen Zahl von Funktionen anzuführen, mag es nützlich sein, verschiedene Möglichkeiten ihrer strukturellen Organisation zu betrachten. Klarerweise haben wir es hier mit Aggregaten von großen Zahlen von Maschinen mit endlich vielen Zuständen zu tun, und es bedarf eines leistungsfähigeren Notationssystems, die durch solche Aggregate ausgeführten Operationen abzubilden.

2 *Tesselierungen*

Obwohl eine Maschine mit endlich vielen Zuständen aus drei verschiedenen Teilen besteht, d.h. den zwei Rechnern f_y und f_z und dem Speicher für z (vgl. Bild 3a), möchte ich die ganze Maschine nur durch ein Quadrat (oder Rechteck) darstellen; ihr Inputbereich ist weiß abgebildet, ihr Outputbereich schwarz (Bild 6).

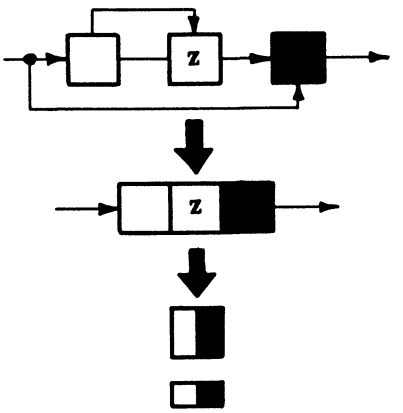


Bild 6
Symbolisierung einer Maschine mit endlich vielen Zuständen durch eine Rechenplatte; Inputbereich weiß, Outputbereich schwarz.

Ich werde diese Einheit nun als einen elementaren Rechner behandeln – als eine „Rechenplatte", T_i –, die durch Kombination mit anderen Platten, T_j, ein Mosaik von Platten bilden kann – eine „Rechentesselierung", J. Die durch die i-te Platte ausgeführten Operationen sollen die einer Maschine mit endlich vielen Zuständen sein, ich werde aber statt Subskripten verschiedene Buchstaben verwenden, um die beiden charakteristischen Funktionen zu unterscheiden. Subskripte sollen sich jeweils auf Platten beziehen.

$$\begin{aligned} y_i &= f_i(x_i, z_i) \\ z_i &= g_i(x_i, z_i). \end{aligned} \tag{48}$$

[4]) Im Falle mehrerer Gleichgewichtszustände, $\{f_{0i}\}$, haben wir natürlich eine *Menge* nicht-trivialer Maschinen mit endlich vielen Zuständen, die das Resultat unterschiedlicher Ausgangsbedingungen darstellen.

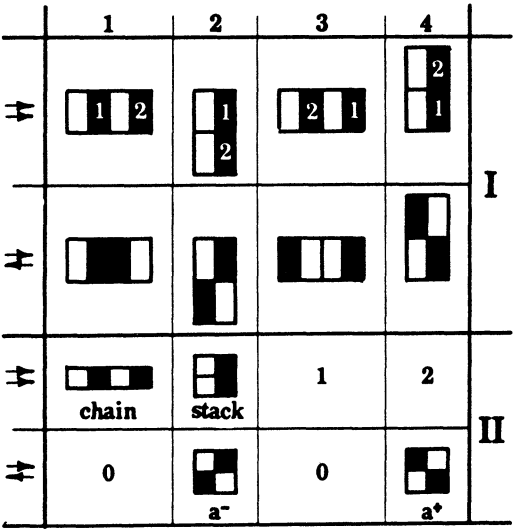


Bild 7
Elementare Tesselierungen

Bild 7/I skizziert die acht möglichen Arten (jeweils vier für den parallelen und den antiparallelen Fall), auf die zwei Platten miteinander verbunden werden können. Daraus ergeben sich zwei Klassen elementarer Tesselierung, deren Strukturen durch Bild 7/II dargestellt werden. Die Fälle I/1 und I/3 sowie I/2 und I/4 sind im Parallelfall äquivalent und werden durch II/1 („Kette") bzw. II/2 („Stapel") repräsentiert. Im antiparallelen Fall sind die beiden Konfigurationen I/1 und I/3 unbrauchbar, denn Outputs können nicht auf Outputs und Inputs nicht auf Inputs wirken; die Fälle I/2 und I/4 erzeugen zwei autonome elementare Tesselierungen $A = [a^+, a^-\}$, die nur der Rotationsrichtung nach, in der die Signale verarbeitet werden, verschieden sind.

Die Iterierung derselben Verkettungen ergibt Tesselierungen mit den folgenden funktionalen Eigenschaften (für n Iterierungen):

1. Stapel $\quad nT: \; y = \sum_{1}^{n} f_i(x_i, z_i)$ (49)

2. Kette $\quad T^n: \; y = f_n(f_{n-1}(f_{n-2} \ldots (x^{(n)*}, z^{(n)*}) \ldots z^{**}_{n-2}) z^{*}_{n-1}) z_n$ (50)

3. $\quad A = \{a^+, a^-\}$

$$\left.\begin{matrix} a^+a^- \\ a^-a^+ \end{matrix}\right\} = 0 \qquad \left.\begin{matrix} a^+a^+ \\ a^-a^- \end{matrix}\right\} \neq 0$$

(i) *Stapel* $\quad nA^n$ (51)

(ii) *Kette* $\quad A^n$ (52)

Führt man eine vierte elementare Tesselierung dadurch ein, daß man $T \rightarrow A \rightarrow T$, oder TAT, verknüpft, dann ergibt sich

4. TAT

(i) *Stapel* $\quad n(TA^nT)$ (53)

(ii) *Kette* $\quad (TAT)^n$ (54)

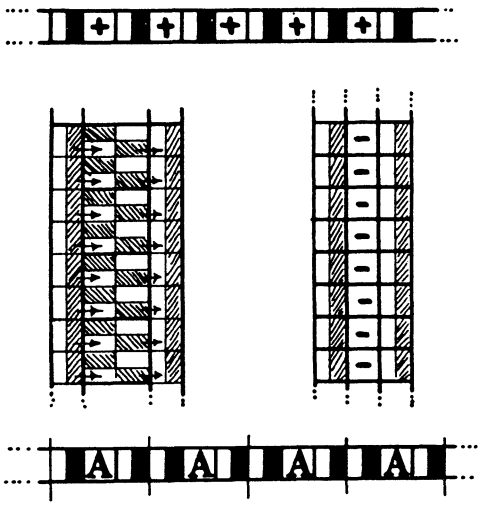


Bild 8
Einige Beispiele für einfache Tesselierungen.

Bild 8 zeigt weitere Gebilde elementarer Tesselierung. Alle diese enthalten autonome Elemente, denn das Auftreten von zumindest zwei solchen Elementen, wie z.B. in $(TAT)^2$, konstituiert eine Maschine mit endlich vielen Funktionen. Ist keines dieser Elemente „tot" – d.h. auf einen einzigen Zustand statischen Gleichgewichts festgelegt –, so zwingen diese Elemente einander, von einem dynamischen Gleichgewicht zu einem anderen überzugehen. Mit anderen Worten, sie verwandeln einander unter bestimmten Umständen aus einer trivialen Maschine mit endlich vielen Funktionen in eine andere, dies ist aber genau das Kriterium dafür, eine nicht-triviale Maschine mit endlich vielen Funktionen zu sein.

Es ist hier darauf hinzuweisen, daß die Vorstellung, daß formale mathematische Entitäten miteinander interagieren, nicht neu ist. John von Neumann (1966) hat dieses Konzept für sich selbst reproduzierende „Automaten" ausgearbeitet, die viele Eigenschaften mit unseren Platten gemeinsam haben. Lars Löfgren (1962) hat dieses Konzept so erweitert, daß es die selbsttätige Reparatur gewisser Rechenelemente einschließt, die entweder stationär sind oder sich in ihren Tesselierungen frei bewegen, und Gordon Pask schließlich hat (1962) ähnliche Ideen entwickelt, um die soziale Selbstorganisation von Aggregaten solcher Automaten erörtern zu können.

Es ist festzuhalten, daß in all diesen Untersuchungen *Mengen* von Elementen betrachtet werden, um bei der Diskussion der Eigen-Begriffe und Autonomie-Eigenschaften der behandelten Elemente, wie z.B. *Selbst*-Replikation, *Selbst*-Reparatur, *Selbst*-Organisation, *Selbst*-Erklärung usw. logische Geschlossenheit zu erreichen. Dies ist kein Zufall, wie Löfgren (1968) beobachtet hat, denn das Präfix „Selbst-" kann durch jenen Term ersetzt werden, dessen Präfix es ist, um einen Begriff zweiter Ordnung zu erzeugen, d.h. den Begriff eines Begriffs. Selbst-Erklärung ist die Erklärung einer Erklärung; Selbst-Organisation ist die Organisation einer Organisation (Selfridge, 1962) etc. Da Kognition im wesentlichen ein selbstreferentieller Prozeß ist (von Foerster, 1969), ist zu erwarten, daß wir bei der Erörterung der ihr zugrundeliegenden Mechanismen Funktionen von Funktionen und Strukturen von Strukturen zu betrachten haben.

Da im Aufbau dieser Strukturen ihre funktionale Komplexität rasch zunimmt, würde eine detaillierte Diskussion ihrer Eigenschaften den Rahmen dieser Arbeit überschreiten. *Ein* Merkmal dieser Rechentesselierungen ist jedoch leicht zu erkennen, und zwar, daß ihre operationalen Modalitäten eng mit ihrer strukturellen Organisation verbunden sind. Hier gehen Funktion und Struktur Hand in Hand, und man sollte nicht übersehen, daß der größte Teil der Rechenarbeit vielleicht bereits geleistet worden ist, wenn die Topologie des Systems vorliegt (Werner 1970). Dies geschieht in Organismen natürlich primär durch genetische Rechenprozesse.

Diese Beobachtung führt uns direkt zur Physiologie und zur Physik organischer Tesselierungen.

III Biophysik

A Allgemeine Bemerkungen

Es stellt sich nun die Frage, ob man strukturelle oder funktionale Einheiten in lebenden Organismen eingrenzen kann, die im Sinne der zuvor erwähnten rein mathematischen Objekte, also der „Platten", „Automaten", „Maschinen mit endlich vielen Funktionen" usw. interpretiert werden können. Diese Verfahrensweise, nämlich zuerst ein Modell anzufertigen und dann nach bestätigenden Daten zu suchen, scheint der „wissenschaftlichen Methode" zu widersprechen, nach der die „Tatsachen" ihrer Erklärung vorausgehen sollen. Was jedoch als „Tatsache" festgestellt wird, ist bereits durch das kognitive System des Beobachters gegangen, welches diesen ja sozusagen mit apriorischen Interpretationen versorgt. Da unser Geschäft hier darin besteht, die Mechanismen zu bestimmen, die Beobachter beobachten (d.h. zu „Selbst-Beobachtern" werden), ist es unser gutes Recht, zuerst die notwendigen funktionalen Strukturen dieser Mechanismen zu postulieren. In der Tat ist dies außerdem ein sehr gängiger Ansatz, wie sich am häufigen Gebrauch von Ausdrücken wie „Spur", „Engramm", „Speicher", „Einlesen", „Ausdrucken" usw. beobachten läßt, die in der Erörterung der Mechanismen des Gedächtnisses auftreten. Auch hier geht klarerweise die Metapher den Beobachtungen voraus. Metaphern haben jedoch mit Interpretationen die Eigenschaft gemein, daß sie weder wahr noch falsch sind: sie sind lediglich entweder nützlich, nutzlos oder irreführend.

Wenn eine funktionale Einheit begrifflich eingegrenzt wird – ein *Lebewesen*, ein *Gehirn*, das *Kleinhirn*, *neuronale Nuklei*, ein *einzelnes Neuron*, eine *Synapse*, eine *Zelle*, die *Organellen*, die *Genome*, und andere molekulare Bausteine –, dann ist die Anwendung des Begriffs der Maschine in seinem abstrakten Sinne auf diese Einheiten nützlich, auch wenn dies nur dazu diente, den Verwender dieses Begriffes so zu disziplinieren, daß er die strukturellen und funktionalen Bestandteile seiner „Maschine" in angemessener Weise bestimmt. In der Tat haben die Konzepte der Maschine mit endlich vielen Zuständen bzw. alle ihre methodologischen Verwandten – explizit oder implizit – erheblich zum Verständnis einer großen Vielfalt derartiger funktionaler Einheiten beigetragen. Die Nützlichkeit etwa der Begriffe „Transkript", „Enkodierung", „Dekodierung", „Rechnen" usw. in der Molekulargenetik ist kaum zu bestreiten.

Sei nun die n-Sequenz der vier Basen (b = 4) eines bestimmten DNS-Moleküls durch eine ν-Zahl $\nu(n, b)$ repräsentiert [vgl. Gleichung (33)]; sei $\mathrm{Tr}(\nu) = \overline{\nu}$ eine

Operation, die die Symbole (0, 1, 2, 3) in eben dieser Ordnung zu (3, 2, 1, ∅) transformiert, wobei 0 ≡ Thymin, 1 ≡ Cytosin, 2 ≡ Guanin, 3 ≡ Adenin und ∅ ≡ Uracil, und sei I die Identitätsoperation $I(\nu) = \nu$; sei schließlich $\Phi[\overline{\nu}(n, b)] = \nu(n/3, a) = \mu(m, a)$, mit $a = 20$ und $j = 0, 1, \ldots, 19$, die Repräsentation der 20 Aminosäuren der Polypeptidkette. Damit ergibt sich

(i) DNS Replikation: $\nu = I(\nu)$ (55a)

(ii) DNS/RNS Transkript: $\overline{\nu} = Tr(\nu)$ (55b)

(iii) Proteinsynthese: $\mu = \Phi(\overline{\nu})$. (55c)

Während die Operationen I und Tr lediglich triviale Maschinen für den Transkriptionsprozeß benötigen, ist Φ eine rekursive Berechnung von der Form

$$j(i) \equiv y(i) = y(i-1) + a^i f(x). \tag{56}$$

Wenden wir nun die vorgeschlagene Rekursion an [vgl. Gleichung (14)]:

$$y(i) = a^i f(x) + a^{i-1} f(x^*) + a^{i-2} f(x^{**}) \ldots$$

oder

$$y(i) = \sum_{k=0}^{i} a^{i-k} f(x^{(k)*}) \tag{57}$$

und

$$y(m) \equiv \mu(m, a).$$

Die Funktion f wird natürlich von dem Ribosom berechnet, welches das Kodon x liest und die Aminosäuren synthetisiert, die sodann durch die Rekursion zu einer verbundenen Polypeptidkette verknüpft werden.

Die Veranschaulichung des ganzen Prozesses durch die Operationen einer sequentiellen Maschine endlich vieler Zustände war wahrscheinlich mehr als nur ein Hinweis für das Gelingen der „Entschlüsselung des genetischen Kodes" und der Identifizierung des Inputzustandes dieser Maschine als Tripel (u, v, w) angrenzenden Symbole in der $\overline{\nu}$-Zahl-Repräsentation der Boten-RNS.

Eine Methode zur Berechnung von ν-Zahlen molekularer Sequenzen direkt aus den Eigenschaften der generierten Struktur wurde von Pattee vorgeschlagen (1961). Er verwendete die Vorstellung eines sequentiellen „Verschubregisters", also das Prinzip einer autonomen Platte. In der Errechnung periodischer Sequenzen in wachsenden gewendelten Molekülen wird das Element, das als jeweils nächstes dem Wendel angefügt werden soll, ausschließlich durch die gegebenen und einige frühere Bausteine determiniert. Es bedarf keinerlei externen Rechensystems.

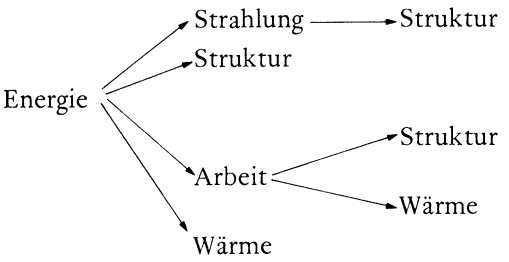
Wenn man auf einer höheren Ebene der hierarchischen Organisation das Neuron als funktionale Einheit annimmt, dann lassen sich viele Beispiele geben, in denen dieses als Errechner rekursiver Funktionen betrachtet werden kann. Je nachdem, was als „Signal" angesehen wird, ein einziger Puls, ein Kode von durchschnittlicher Häufigkeit, ein Latenzkode, ein Wahrscheinlichkeitskode (Bullock 1968) usw., wird das Neuron zu einem „Alles-oder-Nichts"-Instrument der Berechnung logischer Funktionen (McCulloch/Pitts 1943), ein lineares Element

(Sherrington 1906), ein logarithmisches Element etc., wobei sich eigentlich nur ein einziger für dieses Neuron charakteristischer Parameter verändert (von Foerster 1967b). Das gleiche gilt für Nervennetze, in denen diese Rekursion durch Schleifen oder manchmal direkt durch rekurrente Fasern hergestellt wird. Das „widerhallende" neuronale Netz ist ein typisches Beispiel für eine Maschine mit endlich vielen Zuständen im dynamischen Gleichgewicht.

Angesichts wohl einer ganzen Bibliothek voll der Beispiele, in denen der Begriff der Maschine mit endlich vielen Zuständen sich als nützlich erwiesen hat, mag es nun überraschend wirken, wenn man feststellt, daß derartige Systeme rein physikalisch völlig absurd sind. Um in Gang zu bleiben, müssen sie *perpetua mobilia* sein. Dies kann zwar von einem mathematischen Objekt ohne Schwierigkeiten geleistet werden, ist aber für ein reales Objekt ausgeschlossen. Natürlich ist es von einem heuristischen Standpunkt aus irrelevant, ob ein Modell physikalisch verwirklichbar ist oder nicht, solange es in sich konsistent ist und unseren Geist zu weiteren Untersuchungen anspornt.

Wenn jedoch der Energiefluß zwischen verschiedenen Ebenen der Organisation vernachlässigt und die Mechanismen der Energieumwandlung und der Energieübertragung ignoriert werden, dann ergeben sich Schwierigkeiten bei der Zuordnung der deskriptiven Parameter funktionaler Einheiten auf einer Ebene zu solchen auf höheren oder niedrigeren Ebenen. Eine Beziehung etwa zwischen dem Kode eines bestimmten RNS-Moleküls und, sagen wir, dem Pulsfrequenzkode im gleichen Neuron läßt sich nicht herstellen, wenn die Mechanismen der Energieübertragung nicht berücksichtigt werden. Solange die Frage nicht gestellt wird, was den Organismus in Gang hält und wie dies geschieht, bleibt die Kluft zwischen funktionalen Einheiten auf verschiedenen Ebenen der Organisation unüberbrückt. Läßt sich dies mit Hilfe der Thermodynamik durchführen?

Im folgenden sollen drei verschiedene Arten molekularer Mechanismen kurz erörtert werden, die sich für diese Zwecke unmittelbar anbieten. Alle diese gebrauchen verschiedene Formen der Energie, als Strahlung (νh), potentielle Energie (V, Struktur), Arbeit ($p \Delta v$), und Wärme ($k \Delta t$), und deren verschiedene Konversionen von einer Form in eine andere.



Wir verbleiben damit in der Terminologie der Maschine mit endlich vielen Zuständen und klassifizieren die drei Arten des Mechanismus nach ihren Inputs und ihren Outputs, vernachlässigen jedoch für den Augenblick alle Unterscheidungen von Energieformen, abgesehen von der potentiellen Energie (Struktur), die von allen anderen Formen (Energie) verschieden ist.

(i) Molekularspeicher: Energie hinein,
Energie heraus.

(ii) Molekularrechner: Energie hinein,
Struktur heraus.

(iii) Molekularträger: Energie und Struktur hinein,
Energie heraus.

Diese drei Fälle sollen nun kurz erörtert werden.

B Molekularspeicher

Der wahrscheinlich naheliegendste und daher vielleicht älteste Ansatz zur Verknüpfung makroskopischen Verhaltens, wie z.B. des Vergessens von Unsinnssilben (Ebbinghaus 1885), mit dem quantenmechanischen Abklingen der zahlreichen erregten, metastabilen Zustände von Makromolekülen, nimmt weiter nicht analysierbare „Elementareindrücke" an, die mit dem metastabilen Zustand eines Moleküls verknüpft werden (von Foerster 1948; von Foerster 1949). Diese können durch eine nicht-destruktive Ablesung auf ein anderes Molekül übertragen werden, und eine Aufzeichnung dieser Elementareindrücke kann entweder zerfallen oder aber wachsen, je nachdem, ob das Produkt aus Quantenzerfallszeitkonstante und Ablesegeschwindigkeit größer oder kleiner ist als 1. Auch wenn dieses Modell eine gute Übereinstimmung zwischen makroskopischen Variablen, z.B. den Vergessensraten, oder der Temperaturabhängigkeit von vorgestellten Zeitsprüngen (Hoagland 1951; Hoagland 1954), und mikroskopischen Variablen wie Bindeenergien oder Elektron-Orbitalsequenzen herzustellen erlaubt, leidet es wie alle Aufzeichnungsmodelle daran, daß es völlig außerstande ist, aus den angesammelten Aufzeichnungen irgendetwas an Folgerungen abzuleiten. Nur dann, wenn dieser Sammlung eine induktive Inferenzmaschine, die die angemessenen Verhaltensfunktionen berechnet, angeschlossen wird, kann ein Organismus überleben (von Foerster u.a. 1968). Man kann daher alle Spekulationen über Systeme, die Merkmale aufzeichnen, sein lassen und sich jenen zuwenden, die Verallgemeinerungen berechnen.

C Molekularrechner

Die gute Übereinstimmung zwischen makroskopischen und mikroskopischen Variablen des eben besprochenen Modells legt nahe, daß diese Beziehung weiter verfolgt werden sollte. Es läßt sich in der Tat zeigen (von Foerster 1969), daß die Energieintervalle zwischen erregten metastabilen Zuständen so organisiert sind, daß die Abklingzeiten im Gitterschwingungsband neuronalen Pulsintervallen, und ihre Energieebenen einem Polarisationspotential von 60 mV bis 150 mV entsprechen. Eine Pulsfolge von verschiedenen Pulsintervallen „pumpt" daher ein solches Molekül in höhere Erregungszustände, die von den Ausgangsbedingungen abhängen. Erreicht die Erregungsebene jedoch etwa 1,2 eV, dann erfährt das Molekül strukturelle Veränderungen, die jeweils einen Tag oder länger anhalten können. In diesem „strukturell geladenen" Zustand kann es nun auf verschiedene Arten an der Veränderung der Transferfunktionen eines Neurons mitwirken, indem es seine Energie entweder auf andere Moleküle überträgt oder deren Reaktionen erleichtert. Da in diesem Modell ungerichtete potentielle elektrische Energie eingesetzt wird, um spezifische strukturelle Veränderungen herbeizuführen, wird es mit

„Energie hinein – Struktur heraus" bezeichnet. Dies führt zur Vorstellung einer molekularen Errechnung, deren Ergebnis die Ablagerung von Energie an einem spezifischen Verwendungsort ist. Damit beschäftigt sich das nächste und letzte Modell.

D Molekularträger

Eines der am häufigsten verwendeten Prinzipien der Energiedissemination in einem lebenden Organismus liegt in der Trennung des Orts der Synthese von dem der Verwendung. Die allgemein für den Transfer verwendete Methode besteht in einer zyklischen Operation, die einen oder mehrere molekulare Träger benutzt, welche dort, wo Umweltenergie absorbiert werden kann, „aufgeladen", und dort „entladen" werden, wo diese Energie eingesetzt werden muß. Aufladen und Entladen wird gewöhnlich mit Hilfe chemischer Modifikationen der basalen Trägermoleküle bewerkstelligt. Ein einleuchtendes Beispiel für den gerichteten Fluß der Energie und den zyklischen Fluß der Materie ist natürlich die Komplementarität der Prozesse der Photosynthese und der Atmung (Bild 9).

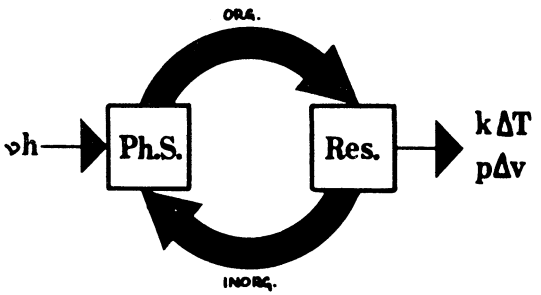


Bild 9
Gerichteter Fluß der Energie und zyklischer Fluß der Materie in der Kopplung von Photosynthese (Ph.S) und Atmung (Res.).

Lichtenergie, νh, bricht die stabilen Bindungen der anorganischen Oxyde und transformiert diese in energetisch aufgeladene organische Moleküle. Diese werden sodann im Atmungsprozeß verbrannt und geben ihre Energie als Arbeit, $p\Delta v$, oder Wärme, $k\Delta T$, am Ort ihrer Verwendung ab, und kehren darauf als anorganische Oxyde wieder an den Ort der Synthese zurück.

Ein weiteres Beispiel ist der äußerst komplizierte Prozeß der Energiemobilisierung in den Mitochondrien. Diese Reaktion synthetisiert nicht nur Adenosintriphosphat (ATP), indem Adenosindiphosphat (ADT) mit einer Phosphatgruppe gekoppelt wird, sondern lädt auch das ATP-Molekül mit beträchtlicher Energie auf, die bei der Muskelkontraktion sehr wirksam wieder abgegeben wird; der Kontraktionsprozeß konvertiert ATP wieder zu ADP, indem die zuvor angebundene Phosphatgruppe abgespalten wird.

Schließlich sei noch die Boten-RNS als Beispiel für getrennte Orte der Synthese bzw. Verwendung zitiert, obwohl die Energetik dieses Falls bis jetzt noch nicht so gut bekannt ist wie in den anderen Fällen. Auch hier ist es offensichtlich die Struktur, die von einem Ort zum anderen übertragen werden muß, und nicht die Energie.

All diesen Prozessen gemeinsam ist die Tatsache, daß im Prozeß der Synthese nicht nur ein wieder abrufbares Quantum Energie, ΔE, auf den molekularen Träger geladen wird, sondern auch eine Adresse, die angibt, wo das Paket abgeladen werden soll. Diese Adresse erfordert einen zusätzlichen Organisationsbetrag, $-\Delta H$, (Negentropie), um ihren Bestimmungsort zu lokalisieren. Es ergibt sich damit die entscheidende Bedingung

$$\frac{\Delta E}{\Delta H} < 0, \tag{58}$$

die besagt: „Bei hoher Energie ist niedrige Entropie, bei niedriger Energie hohe Entropie erforderlich". Dies widerspricht natürlich dem gewöhnlichen Verlauf der Ereignisse, wo diese beiden Quantitäten in einer positiven Beziehung verknüpft sind.

Es läßt sich jedoch zeigen, daß dann, wenn ein System aus Bestandteilen aufgebaut ist, die im Basiszustand getrennt sind, die jedoch, wenn „erregt", in „vernünftig stabilen" metastabilen Zuständen zusammenhängen, die oben formulierte entscheidende Bedingung erfüllt ist (von Foerster 1964).

Sei

$$V = \pm \left(Ae^{-x/\kappa} + B \sin\frac{2\pi x}{p}\right) \tag{59}$$

mit

$$A/B \gg 1 \quad \text{und} \quad \kappa/p \gg 1$$

die Potentialverteilung in zwei eindimensionalen, linearen, „periodischen Kristallen" C^+ und C^-, wobei das $\pm$ sich auf entsprechende Fälle bezieht. Der wesentliche Unterschied zwischen diesen beiden linearen Strukturen, die sich als lineare Verteilungen elektrischer Ladungen vorstellen lassen, die ihr Vorzeichen (fast) periodisch ändern, besteht darin, daß Energie benötigt wird, um den „Kristall" C^+ zusammenzusetzen, und daß es etwa dieselbe Energie braucht, um den Kristall C^- in seine Bestandteile zu zerlegen. Diese linearen Gitter haben metastabile Gleichgewichtszustände bei

$$C^+ \rightarrow x_1, x_3, x_5 \ldots$$
$$C^- \rightarrow x_0, x_2, x_4 \ldots,$$

die Lösungen sind von

$$e^{x/\kappa} \cos\frac{2\pi x}{p} = \frac{1}{2\pi}\frac{Ap}{B\kappa} \approx 1.$$

Diese Zustände sind durch eine Energieschwelle geschützt, die sie für den durchschnittlichen Zeitbetrag

$$\tau = \tau_0 e^{\Delta\nu/kT} \tag{60}$$

in diesem Zustand erhält, wobei τ_0^{-1} die Elektron-Orbital-Frequenz ist und ΔV die Differenz zwischen den Energien im Tal und am Kamm der Potentialwelle $[\pm \Delta V_n = V(x_n) - V(x_{n+1})]$.

Um die Entropie dieser Konfiguration festzustellen, lösen wir die Schrödinger-Gleichung (die hier in normalisierter Form gegeben wird)

$$\psi'' + \psi[\lambda - V(x)] = 0 \tag{61}$$

für ihre Eigenwerte λ_i und ihre Eigenfunktionen ψ_i, ψ_i^*, die sodann die Wahrscheinlichkeitsverteilung dafür liefern, daß das Molekül im i-ten Eigenzustand ist:

$$\left(\frac{dp}{dx}\right)_i = \psi_i \cdot \psi_i^*, \tag{62}$$

natürlich mit

$$\int_{-\infty}^{+\infty} \psi_i \cdot \psi_i^* \, dx = 1, \tag{63}$$

woraus wir die Entropie

$$H_i = -\int_{-\infty}^{+\infty} \psi_i \cdot \psi_i^* \ln \psi_i \cdot \psi_i^* \tag{64}$$

für den i-ten Eigenzustand ableiten.

Es ist bemerkenswert, daß die Veränderung des Energie-Entropieverhältnisses für die beiden Kristalle C^+ und C^- aufgrund der Aufladung ($\Delta E = e(V(x_n) - V(x_{n+2}))$ in zwei entgegengesetzten Richtungen verläuft:

$$C^- \rightarrow \left(\frac{\Delta E}{\Delta H}\right)^- > 0$$

$$C^+ \rightarrow \left(\frac{\Delta E}{\Delta H}\right)^+ < 0.$$

Damit ist gezeigt, daß die beiden Kristalle ganz verschiedene Lebewesen sind: der eine ist tot (C^-), der andere lebendig (C^+).

IV Zusammenfassung

Diese Arbeit ist im Kern ein Versuch, die ursprüngliche Bedeutung von Begriffen wie Gedächtnis, Lernen, Verhalten usw. wieder herzustellen, und zwar dadurch, daß sie als unterschiedliche Manifestationen des umfassenderen Phänomens der Kognition aufgefaßt werden. Es wird der Versuch gemacht, diese Behauptung zu rechtfertigen und einen Begriffsapparat von genügender Reichhaltigkeit zu skizzieren, um diese Phänomene in ihrem angemessenen Umfang zu beschreiben. Dieser Vorschlag wurde in seiner prägnantesten Form als eine Suche nach Mechanismen in lebenden Organismen dargestellt, die diese instandsetzen, ihre Umwelt in eine triviale Maschine zu verwandeln, und nicht so sehr als die Suche nach Mechanismen der Umwelt, die die Organismen in triviale Maschinen verwandeln.

Dieser Standpunkt wird durch die Erkenntnis gerechtfertigt, daß der zweite Ansatz – wenn er Erfolg hat – keine Erklärung für die Mechanismen liefert, die er zu entdecken sucht, denn eine triviale Maschine zeigt keine der gewünschten

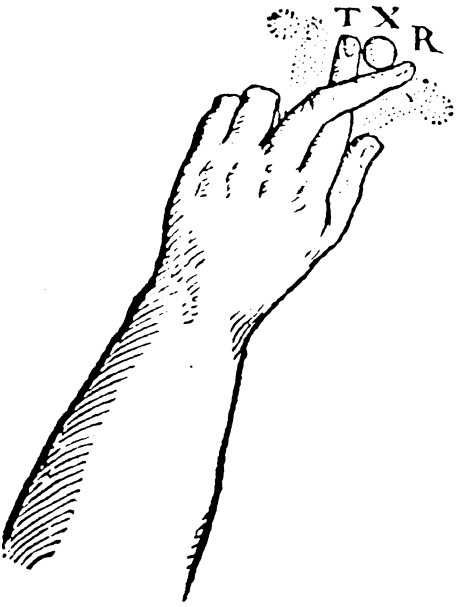
Eigenschaften; und wenn er scheitert, dann enthüllt er nicht die Ursachen für dieses Scheitern.

Im Rahmen des Begriffssystems der Maschinen mit endlich vielen Zuständen wurde der Kalkül rekursiver Funktionale als ein deskriptiver (phänomenologischer) Formalismus vorgeschlagen, um Gedächtnis als potentielle Vergegenwärtigung früherer Interpretationen von Erfahrungen zu begründen, folglich auch, um den Ursprung des *Begriffs* der „Veränderung" und die Übergänge zwischen Bereichen zu erklären, wie sie dann auftreten, wenn wir von „Tatsachen" zu „Beschreibungen von Tatsachen" und – da diese ihrerseits Tatsachen sind – zu „Beschreibungen von Beschreibungen von Tatsachen" kommen usw.

Elementare Maschinen mit endlich vielen Funktionen lassen sich verknüpfen, um lineare oder zweidimensionale Tesselierungen von beträchtlicher rechnerischer Flexibilität und Komplexität zu bilden. Derartige Tesselierungen sind nützliche Modelle für Aggregate interagierender funktionaler Einheiten auf verschiedenen Ebenen der hierarchischen Organisation von Organismen. Auf der molekularen Ebene z.B. kann eine kettenähnliche Tesselierung, die zu einer Helix gedreht wird, sich selbst (Selbstreplikation) oder in Verbindung mit anderen Elementen weitere funktionale molekulare Einheiten errechnen (Synthese).

Während der Begriff des rekursiven Funktionals in der Erörterung deskriptiver Formalismen die Brücke dafür liefert, verschiedene deskriptive Bereiche zu durchschreiten, verbindet der Begriff der Energieübertragung zusammen mit dem der entropischen Veränderung die funktionalen Einheiten auf verschiedenen Organisationsebenen in operationaler Weise. Eben diese Bindeglieder, seien sie begriffliche oder operationale, schaffen die Voraussetzungen für die Erklärung von Strukturen und Funktionen lebender Organismen, die als autonome selbstreferentielle Organismen angesehen werden. Werden diese Bindeglieder mißachtet, ist der Begriff des Organismus leer, und seine unverbundenen Einzelteile werden entweder zu Trivialitäten oder bleiben Rätsel.

Gegenstände: greifbare Symbole für (Eigen-)Verhalten


T X
R

Gegenstände: greifbare Symbole für (Eigen-)Verhalten *

Ein Keim nur, leider keine Blüte, für Jean Piaget zum 80. Geburtstag in Bewunderung und Zuneigung.

Ich möchte über Begriffe sprechen, die sich ergeben, wenn die Organisation sensorisch-motorischer Interaktionen (und auch zentraler Prozesse (kortikaler-zerebellarer-spinaler, kortikaler-thalamischer-spinaler usw.)) als zirkuläre (oder präziser: rekursive) Organisation aufgefaßt wird. Rekursion spielt in solchen Überlegungen immer dann eine Rolle, wenn die Veränderungen der Sinneswahrnehmungen eines Lebewesens durch dessen Bewegungen ($s_i = S(m_k)$) und seine Bewegungen durch seine Sinneswahrnehmungen bestimmt werden ($m_k = M(s_j)$). Wenn diese beiden Bestimmungen zusammengenommen werden, dann bilden sie „rekursive Ausdrücke", d.h. Ausdrücke, die die Zustände (Bewegungen, Sinneswahrnehmungen) des Systems (des Lebewesens) durch eben diese Zustände selbst festlegen ($s_i = S(M(s_j)) = SM(s_j)$; $m_k = M(S(m_l)) = MS(m_l)$).

Ein Kernpunkt der Überlegungen, der mit mehr Zeit, mehr Aufwand und mehr Raum in exakter Weise – und nicht lediglich andeutungsweise wie hier – dargelegt werden könnte, besteht darin, daß das, was in einer beobachterlosen (linearen, offenen) Epistemologie als „Gegen-Stand" angesehen wird, in einer den Beobachter einbeziehenden (zirkulären, geschlossenen) Epistemologie als „Zeichen für stabile Verhaltensweisen" (oder, wenn man die Terminologie der Theorie rekursiver Funktionen verwendet, als (be-)greifbares „Zeichen für Eigenverhalten") erscheint.

Unter den vielen möglichen Zugängen zu diesem Thema scheint mir der für diese Gelegenheit passendste in dem (rekursiven) Ausdruck zu liegen, der die letzte Zeile auf Seite 63 in Jean Piagets *L'Équilibration des Structures Cognitives* (1975) bildet:

$$Obs.O \rightarrow Obs.S \rightarrow Coord.S \rightarrow Coord.O \rightarrow Obs.O \rightarrow \text{etc.}$$

Dies stellt die Abbildung einer Interaktion zwischen einem Subjekt S und einem Objekt (oder einer Menge von Objekten) O durch einen Beobachter dar. Die in diesem Ausdruck verwendeten Symbole (im zitierten Werk auf Seite 59 definiert) bedeuten (vgl. auch Bild 1):

Obs.S.: „Observable bezüglich der Handlung des Subjekts"
Obs.O.: „Observable bezüglich der Objekte"

---

* Dieser Beitrag wurde ursprünglich aus Anlaß des 80. Geburtstags von Jean Piaget angefertigt und an der Universität Genf am 29. Juni 1976 vorgetragen. Die französische Fassung erschien in: *Hommage à Jean Piaget: Epistémologie génétique et équilibration*, hrsg. von B. Inhelder/R. Garcia/J. Voneche, Neuchatel 1977.

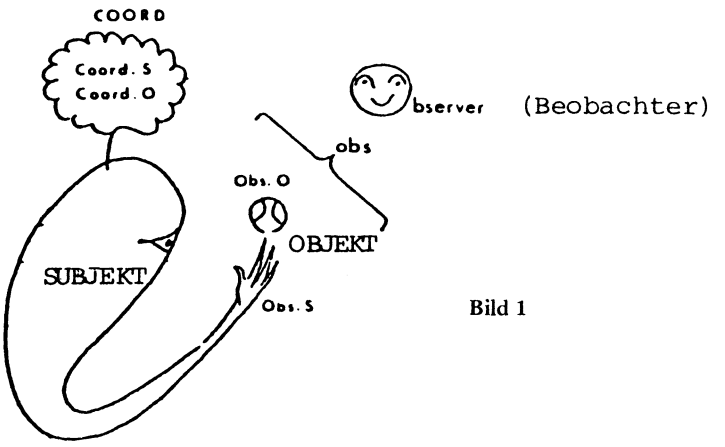


Bild 1

Coord.S.: „Inferentielle Koordinationen der Handlungen (oder Operationen) des Subjekts"

Coord.O.: „Inferentielle Koordinationen zwischen Objekten"

etc.: „der (syntaktische) Befehl, die Sequenz dieser Operationen (ohne festgelegte Grenzen) zu iterieren" (H. v. Foerster).

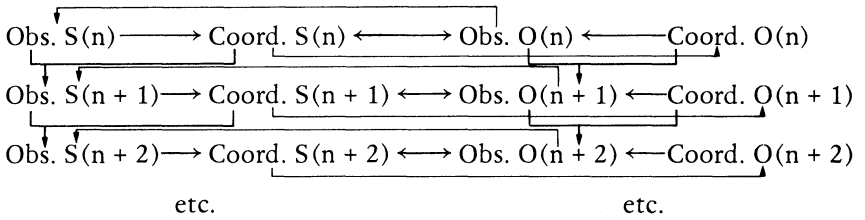
Der Kürze (oder Klarheit?) halber schlage ich vor, den eben zitierten Symbolismus noch weiter zu komprimieren, d.h. alles, was beobachtet wird (d.h. Obs.O und Obs.S) in eine einzige Variable, obs, und die koordinierenden Operationen, die vom Subjekt ausgeführt werden (d.h. Coord.S und Coord.O), in einen einzigen Operator, COORD, zusammenzufassen. COORD transformiert, rearrangiert, modifiziert usw. die Formen, Anordnungen, Verhaltensweisen usw., die in einer bestimmten Situation beobachtet werden (diese sei obs_0 und heiße das „primäre Argument"), zu all jenen, die in der nächsten Situation, obs_1, beobachtet werden. Das Ergebnis dieser Operation sei durch die folgende Gleichung ausgedrückt[1]):

$$obs_1 = \text{COORD}\,(obs_0). \qquad (1)$$

Auch wenn nun in dieser Verdichtung bestimmte relationale Feinstrukturen (klarerweise) verloren gehen, ermöglicht sie einen leichteren Zugang zu der Ab-

[1]) Wenn man den Pfeil „→", dessen operationale Bedeutung im wesentlichen darin liegt, eine einsinnige (semantische) Verknüpfung zwischen nebeneinander stehenden Ausdrücken anzuzeigen (z.B. „geht über in", „impliziert", „löst aus", „führt zu" usw.), durch ein Gleichheitszeichen ersetzt, wird die Basis für einen Kalkül geschaffen. Damit dieses Zeichen jedoch in zulässiger Weise benutzt werden kann, müssen die Variablen „obs_i" zum gleichen Bereich gehören. Die Wahl des Bereiches bleibt natürlich dem Beobachter überlassen. Er kann seine Beobachtungen etwa in numerischen Werten oder in Vektoren als Repräsentationen von Anordnungen oder geometrischen Konfigurationen ausdrücken, er kann im besonderen seine Beobachtungen von Verhaltensweisen durch mathematische Funktionen (z.B. Bewegungsgleichungen usw.) oder logische Propositionen (z.B. McCulloch-Pitts' 1943 „TPEs" (d.s. temporal-propositionale Ausdrücke) usw.) darstellen.

folge von Ereignissen, wie sie in den letzten Zeilen auf Seite 62 des zitierten Werkes formuliert und nachstehend wiedergegeben wird:



Der Operator COORD werde nun auf das zuvor erhaltene Ergebnis angewandt; es ergibt sich

$$obs_2 = COORD(obs_1) = COORD(COORD(obs_0)) \tag{2}$$

und (rekursiv) nach n Schritten

$$obs_n = \underbrace{COORD(COORD(COORD(\ldots}_{\text{n-mal}}\underbrace{(obs_0)))\ldots)}_{\text{n-mal}}, \tag{3}$$

oder in verkürzter Notation

$$obs_n = COORD^{(n)}\,(obs_0). \tag{4}$$

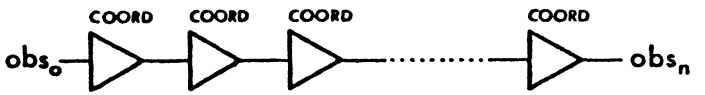


Bild 2

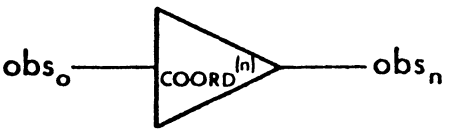


Bild 3

Diese Abkürzung der Notation zeigt außerdem, daß funktional Bild 2 durch Bild 3 ersetzt werden kann.

* * * * *

Möge nun n ohne Begrenzung ($n \to \infty$) wachsen:

$$obs_\infty = \lim_{n \to \infty} COORD^{(n)}\,(obs_0) \tag{5}$$

oder:

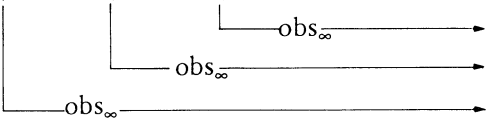
$$obs_\infty = COORD\,(COORD\,(COORD\,(COORD\,\ldots \tag{6}$$

Zum Ausdruck (6) ist zu bemerken:

1. Die unabhängige Variable obs_0, das „primäre Argument", ist verschwunden. (Dies läßt sich als Anzeichen dafür auffassen, daß die einfache Verbindung zwischen unabhängigen und abhängigen Variablen in indefiniten Rekursionen verloren geht, und daß derartige Ausdrücke eine andere Bedeutung annehmen.)

2. Da obs_∞ eine indefinite Rekursion der Operatoren COORD über den Operatoren COORD ausdrückt, kann jede indefinite Rekursion innerhalb dieses Ausdruckes durch obs_∞ ersetzt werden:

obs_∞ =
COORD (COORD (COORD (COORD (...



3. Daraus ergibt sich

$$obs_\infty = obs_\infty \tag{7.0}$$

$$obs_\infty = \text{COORD}(obs_\infty) \tag{7.1}$$

$$obs_\infty = \text{COORD}(\text{COORD}(obs_\infty)) \tag{7.2}$$

$$obs_\infty = \text{COORD}(\text{COORD}(\text{COORD}(obs_\infty))) \tag{7.3}$$

etc.

Obwohl in dieser Darstellung der *horror infinitatis* des Ausdrucks (6) verschwunden ist (alle Ausdrücke in COORD sind finit), so ist doch ein neues Merkmal aufgetreten, nämlich jenes, daß die abhängige Variable obs_∞ sozusagen „von sich selbst abhängig" (oder: „selbstdefinierend", „selbstreflektierend" etc. durch den Operator COORD) geworden ist.

Sollte es nun Werte $obs_{\infty\,i}$ geben, die die Gleichung (7) erfüllen, so mögen diese Werte

„Eigen-Werte",
$obs_{\infty\,i} \equiv Obs_i$ (8)

(oder: „Eigenfunktionen", „Eigenoperatoren", „Eigenalgorithmen", „Eigenverhalten" usw. in Abhängigkeit von dem Bereich der obs) heißen. Sie seien dadurch gekennzeichnet, daß ihr erster Buchstabe großgeschrieben wird. (Beispiele vgl. Anhang A.)

Wir wollen nun Ausdrücke von der Form (7) betrachten und festhalten:

1. Eigenwerte sind diskret (auch wenn der Bereich des primären Arguments obs_0 kontinuierlich ist).

Dies verhält sich so, weil jede infinitesimale Störung $\pm \epsilon$ durch einen Eigenwert Obs_i (das heißt $Obs_i \pm \epsilon$) verschwindet, wie alle anderen Werte von obs verschwunden sind, ausgenommen jene, für die $obs = Obs_i$, und weil obs entweder auf Obs_i (*stabiler* Eigenwert) oder auf einen anderen Eigenwert Obs_j (*instabiler* Eigenwert Obs_i) zurückgeführt wird.

Mit anderen Worten, Eigenwerte stellen *Gleichgewichtszustände* dar. Diese Gleichgewichtszustände können in Abhängigkeit vom gewählten Bereich des primären Arguments Gleichgewichtswerte („Fixpunkte"), funktionale Gleichgewichtszustände, operationale Gleichgewichtszustände, strukturelle Gleichgewichtszustände usw. sein.

2. Die Eigenwerte Obs_i und die ihnen entsprechenden Operatoren COORD stehen in einem komplementären Verhältnis zueinander, d.h. sie implizieren einander; die Obs_i repräsentieren die in der Außenwelt beobachtbaren Manifestationen der (introspektiv zugänglichen) kognitiven Rechenprozesse (Operationen) COORD.

3. Eigenwerte erzeugen aufgrund ihrer selbstdefinierenden (oder selbsterzeugenden) Natur topologische „Geschlossenheit" („Zirkularität") (vgl. Bild 4).

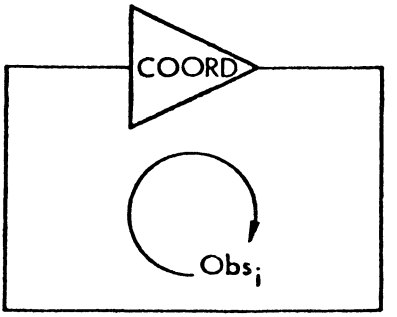


Bild 4

Dieser Sachverhalt erlaubt nun eine symbolische Neuformulierung des Ausdrucks (5):

$$\lim_{n \to \infty} \text{COORD}^{(n)} \equiv \text{COORD}$$

Und dies ist gleichbedeutend mit der Schlange, die sich in den eigenen Schwanz beißt (Bild 5):

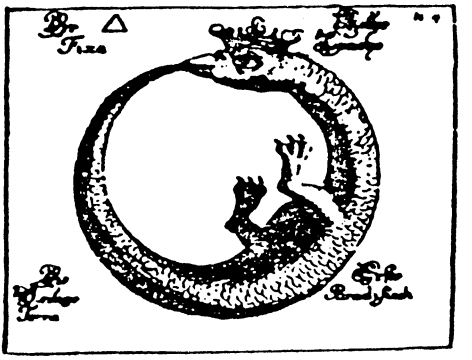


Bild 5

Kognition errechnet ihre eigenen Kognitionen.

* * * * *

Wir wollen nun für einen gegebenen Operator COORD zumindest drei Eigenwerte

$$Obs_1, Obs_2, Obs_3$$

und eine (algebraische) Verknüpfung „*" annehmen, so daß

$$Obs_1 * Obs_2 = Obs_3. \quad (10)$$

Die koordinierenden Operationen COORD koordinieren damit das Ganze (das heißt die Verknüpfung der Teile) als die Verknüpfung der gegebenen Koordinationen der Teile (vgl. Beweis im Anhang B):

$$COORD(Obs_1 * Obs_2) = COORD(Obs_1) * COORD(Obs_2). \quad (11)$$

Mit anderen Worten, die Koordination der Verknüpfungen (d.h. des Ganzen) entspricht der Verknüpfung der Koordinationen.

Dies ist nun die Bedingung für das, was als „Prinzip der kognitiven Kontinuität" bezeichnet werden kann (z.B. ergibt das Auseinanderbrechen von Kreidestücken wiederum Kreidestücke).

Dies läßt sich dem „Prinzip der kognitiven Diversität" gegenüberstellen, das gilt, wenn die Obs_i und die Verknüpfung „*" *nicht* die Eigenwerte und Verknüpfungen sind, die die Koordination COORD′ komplementieren:

$$COORD'(Obs_1 * Obs_2) \neq COORD'(Obs_1) * COORD'(Obs_2). \quad (12)$$

Dieses Prinzip besagt, daß das Ganze zwar nicht mehr, noch aber auch weniger ist als die Summe seiner Teile: das Ganze ist *anders*. Darüber hinaus läßt der Formalismus, in dem diese Auffassung vorgelegt wird (Ausdruck 12), geringen Zweifel daran, daß er weder von „Ganz(heit)en" noch von „Teilen" spricht, sondern von der Unterscheidung, die ein Subjekt zwischen zwei Sachverhalten trifft, welche von einem (anderen) Beobachter nicht qualitativ, sondern lediglich quantitativ geschieden werden.

* * * * *

Es hat sich nun gezeigt, daß Eigenwerte ontologisch diskret, stabil, voneinander trennbar und miteinander verknüpfbar sind, während sie ontogenetisch als Gleichgewichtszustände entstehen, die sich in zirkulären Prozessen selbst bestimmen. Ontologisch können Eigenwerte und Objekte – und entsprechend ontogenetisch: stabiles Verhalten und die Manifestation des „Begreifens" eines Objekts durch ein Subjekt – nicht unterschieden werden. In beiden Fällen sind „Objekte" ausschließlich in die Erfahrung der eigenen sensumotorischen Koordinationen eines Subjekts eingeschlossen, d.h. „Objekte" sind durchwegs subjektiv! Unter welchen Bedingungen erlangen Objekte dann „Objektivität"?

Offensichtlich geschieht dies erst dann, wenn ein Subjekt S_1 die Existenz eines weiteren Subjekts S_2 feststellt, das ihm selbst nicht unähnlich ist, welches seinerseits die Existenz eines weiteren Subjekts, das ihm nicht unähnlich ist, behauptet, das mit S_1 identisch sein kann.

In diesem atomaren sozialen Kontext kann nunmehr die Erfahrung der eigenen sensorisch-motorischen Koordinationen jedes Subjekts (jedes Beobachters)

durch ein Zeichen, d.h. ein „Objekt" repräsentiert werden, das gleichzeitig als Zeichen dafür dient, daß der gemeinsame Raum eine Außenwelt bildet.

Damit bin ich zur Topologie der Geschlossenheit zurückgekehrt (Bild 6):

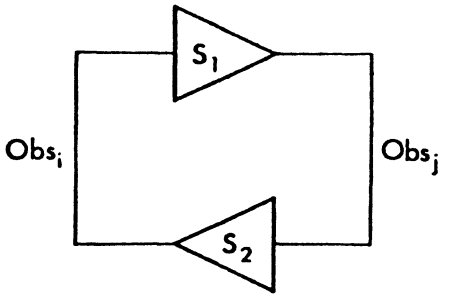


Bild 6

Hier wird Gleichgewicht dann erreicht, wenn das Eigenverhalten eines Beteiligten (rekursiv) das Eigenverhalten eines anderen generiert (vgl. etwa Beispiel A 2 im Anhang); wenn eine Schlange in den Schwanz der anderen Schlange beißt, so als ob es ihr eigener wäre (Bild 7); wenn Kognition ihre eigenen Kognitionen durch die Kognitionen eines anderen errechnet. Hierin liegt der Ursprung der Ethik.

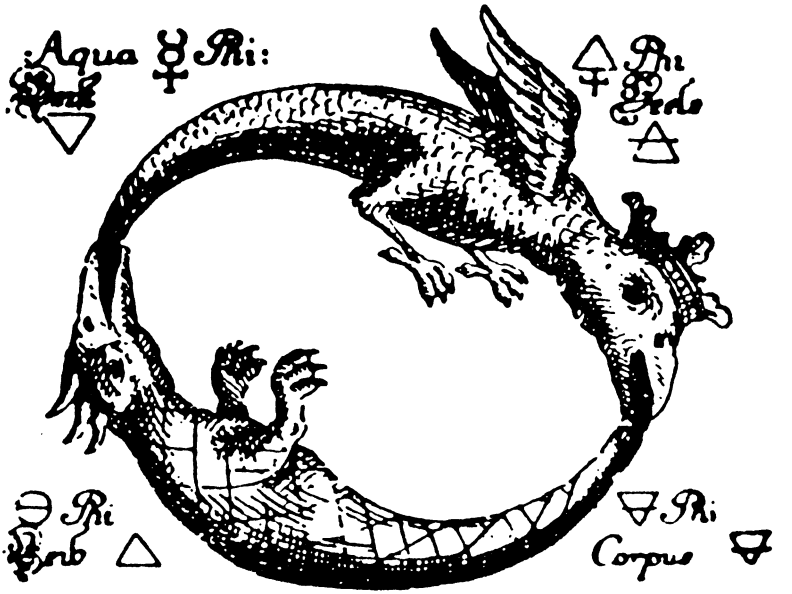


Bild 7

Anhang A

Beispiele

A 1 Nehmen wir den Operator (die lineare Transformation) Op_1,

Op_1 = „Dividiere durch 2 und addiere 1",

und wenden wir ihn (rekursiv) auf x_0, x_1 usw., an (deren Bereiche die reellen Zahlen sind). Wir wählen als Ausgangssymbol x_0 zum Beispiel $x_0 = 4$.

$$x_1 = Op_1(4) = \frac{4}{2} + 1 = 2 + 1 = 3;$$

$x_2 = Op_1(3) = 2.500;$
$x_3 = Op_1(2.500) = 2.250;$
$x_4 = Op_1(2.250) = 2.125;$
$x_5 = Op_1(2.125) = 2.063;$
$x_6 = Op_1(2.063) = 2.031;$
$x_{11} = Op_1(x_{10}) = 2.001;$
$x_\infty = Op_1(x_\infty) = 2.000$

Als weiteren Ausgangswert nehmen wir z.B. $x_0 = 1$.

$x_1 = Op_1(1) = 1.500;$
$x_2 = Op_1(1.500) = 1.750;$
$x_3 = Op_1(1.750) = 1.875;$
$x_8 = Op_1(x_7) = 1.996;$
$x_{10} = Op_1(x_9) = 1.999;$
$x_\infty = Op_1(x_\infty) = 2.000.$

Es ergibt sich in der Tat:

$$\frac{1}{2} \cdot 2 + 1 = 2$$

$Op_1(2) = 2,$

d.h. „2" ist der (einzige) Eigenwert von Op_1.

A 2 Nehmen wir nun den Operator Op_2:

$Op_2 = \exp(\cos\ )$.

Es gibt hier drei Eigenwerte, wovon zwei einander implizieren („Bi-Stabilität") und der dritte instabil ist:

$Op_2(2.4452\ldots) = 0.4643\ldots$
$Op_2(0.4643\ldots) = 2.4452\ldots$ stabil
$Op_2(1.3029\ldots) = 1.3092\ldots$ instabil.

Dies bedeutet:

$Op_2^{(2)}(2.4452\ldots) = 2.4452$ stabil
$Op_2^{(2)}(0.4643\ldots) = 0.4643$ stabil.

A 3 Nehmen wir nun den Differentialoperator Op_3:

$$Op_3 = \frac{d}{dx}.$$

Die Eigenfunktion für diesen Operator ist die Exponentialfunktion „exp":

$$Op_3(exp) = exp,$$

d.h.

$$\frac{de^x}{dx} = e^x.$$

Die Generalisierungen dieses Operators sind natürlich alle Differentialgleichungen, Integralgleichungen, Integral-Differentialgleichungen usw., was sofort einsichtig wird, wenn diese Gleichungen in Operatorform neu geschrieben werden, etwa auf folgende Weise:

$$F(Op_3^{(n)}, Op_3^{(n-1)} \ldots, f) = 0.$$

Natürlich können auch diese Operatoren ihrerseits Eigenwerte (Eigenoperatoren) von „Meta-Operatoren" sein usw. Dies bedeutet, daß COORD zum Beispiel auch selbst als Eigenoperator behandelt werden kann, der innerhalb bestimmter Grenzen stabil ist, aber in andere Werte übergeht, wenn die Grenzbedingungen seinen Stabilitätsbereich verändern:

$$Op(COORD_i) = COORD_i.$$

Man ist schließlich versucht, die Vorstellung eines Meta-Operators zu der eines „Meta-Meta-Operators" zu erweitern, der die „Eigen-Meta-Operatoren" errechnet usw. immer höher in der Hierarchie ohne Ende. Es ist jedoch nicht notwendig, diesen Fluchtweg zu benutzen, wie Warren S. McCulloch schon vor Jahren (1945) in seiner Arbeit „A Heterarchy of Values Determined by the Topology of Nervous Nets" demonstriert hat.

Es würde den Rahmen dieser Darlegung sprengen, hier die Konstruktion von Heterarchien von Operatoren zu demonstrieren, wie sie sich aus deren Verknüpfbarkeit ergeben.

A 4 Nehmen wir die (selbst-referentielle) Aussage

"THIS SENTENCE HAS ... LETTERS"

und vervollständigen wir diese Aussage, indem wir in den freien Raum das entsprechende Wort für die Zahl (oder wenn es mehr als eine gibt, die Zahlen) schreiben, die diese Aussage zu einer wahren Aussage machen.

Wenn wir nach Versuch und Irrtum vorgehen und das, was dieser Satz aussagt (Abszisse), mit dem vergleichen, was er ist (Ordinate), dann finden wir zwei Eigenwerte, nämlich "thirty-one" und "thirty-three". Man wende nun die obige Aussage auf sich selbst an: " 'This sentence has thirty-one letters' has thirty-one letters". Es gilt unter anderem, daß die Aussage "This sentence consists of ... letters" nur *einen* Eigenwert hat (nämlich: thirty-nine), während die Aussage "This sentence is composed of ... letters" *keinen* hat!*

(Fußnote Seite 216)

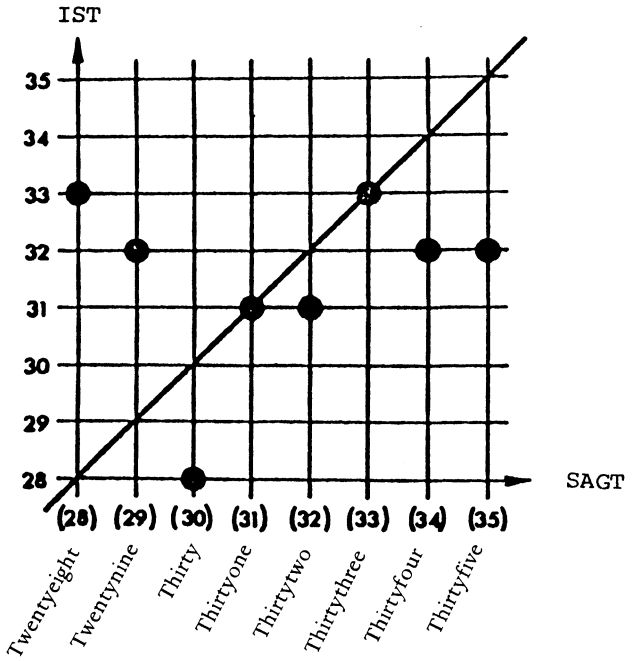


Bild 8

Anhang B

B 1 Beweis des Ausdrucks (11):

$$\text{COORD}(\text{Obs}_1 * \text{Obs}_2) = \text{COORD}(\text{Obs}_3) =$$
$$\text{Obs}_3 = \text{Obs}_1 * \text{Obs}_2 = \text{COORD}(\text{Obs}_1) * \text{COORD}(\text{Obs}_2) \qquad \text{Q.E.D.}$$

Die augenscheinliche Distributivität des Operators COORD über die Verknüpfung „*" sollte nicht so mißverstanden werden, daß „*" eine lineare Verknüpfung bedeutet. So verbinden sich zum Beispiel die Fixpunkte $u_i = \exp(2\pi\lambda i)$ (für $i = 0, 1, 2, 3 \ldots$), die den Operator $\text{Op}(u)$ komplementieren:

$$\text{Op}(u) = u \tan\left(\frac{\pi}{4} \pm \frac{1}{\lambda} \ln u\right) ,$$

in multiplikativer Weise, wobei λ eine beliebige Konstante ist:

$$\text{Op}(u_i \cdot u_j) = \text{Op}(u_i) \cdot \text{Op}(u_j).$$

usw.

Fußnote von Seite 215

* Es sei darauf hingewiesen, daß die deutsche Variante „Dieser Satz hat ... Buchstaben" genau *eine* Lösung hat, d. h. *einen* Eigenwert, deren Ermittlung ich aber dem geneigten Leser überlassen möchte!

Zitierte Literatur

Arnold, P.: "Experiencing the Fourth Spatial Dimension", *Accomplishment Summary 70/71, BCL Report No. 71.2,* The Biological Computer Laboratory, University of Illinois, Urbana/Ill. 1971, 201–216.

Arnold, P.: "A Proposal for a Study of the Mechanisms of Perception of, and Formation of Internal Representations of, the Spatial Fourth Dimension", *Accomplishment Summary 71/72, BCL Report No. 72.2,* The Biological Computer Laboratory, University of Illinois, Urbana/Ill. 1972, 223–236.

Ashby, W. R.: *An Introduction to Cybernetics*, London 1956.

Ashby, W. R.: *The Set Theory of Mechanisms and Homeostasis, Technical Report* 7, Biological Computer Laboratory, University of Illinois, Urbana/Ill. 1962.

Ashby, W. R./Walker, C.: "On Temporal Characteristics of Behavior in Certain Complex Systems", *Kybernetik* 3 (1966), 100–108.

Auerbach, R.: "Organization and Reorganization of Embryonic Cells", in: M. C. Yovits/S. Cameron (eds.), *Self-Organizing Systems,* London 1960, 101–127.

Babcock, M. L. et al.: *Some Principles of Pre-Organization in Self-Organizing Systems,* Technical Report No. 2, Electrical Engineering Research Laboratory, Engineering Experiment Station, University of Illinois, Urbana/Ill. 1960.

Barr, A. H. jr. (ed.): *Fantastic Art, Dada, Surrealism*, 3rd ed., New York 1947.

Bielby, W./Biss, K./Schultz, J./Stahl, F./Woo, T.: "An Answering System for Questions Posed in Natural Language", in: *Cognitive Memory*, Int. Rep. No. II, Coordinated Science Laboratory, University of Illinois, Urbana 1970, 6–45.

Bower, T. G. R.: "The Object in the World of the Infant", *Scientific American* 225/4 (October 1971), 30–38.

Brillouin, L.: *Science and Information Theory,* New York 1962.

Brown, G. S.: *Laws of Form*, New York 1972.

Brün, H.: "Technology and the Composer", in: H. von Foerster (ed.), *Interpersonal Relational Networks,* Cuernavaca: Centro Intercultural de Documentacion 1971, 1–10.

Bullock, T. H.: "Biological Sensors", in: D. L. Arm (ed.), *Vistas in Science*, Albuquerque 1968, 176–206.

Castañeda, C.: *A Separate Reality*, New York 1971.

Dancoff, S. M./Quastler, H.: "The Information Content and Error Rate of Living Things", in: H. Quastler (ed.), *Information Theory in Biology*, Urbana 1953, 263–273.

Defares, J. G./Sneddon, I. N.: *The Mathematics of Medicine and Biology*, Chicago 1961.

Descartes, R.: *L'homme*, Paris 1664. Neudruck in: *Oeuvres de Descartes*, vol. XI (1957), Paris, 119–209.

Descharnes, R.: *Die Welt Salvador Dalis*, Lausanne 1962.

Ebbinghaus, H.: *Über das Gedächtnis. Untersuchungen zur experimentellen Psychologie*, Leipzig 1885.

Eccles, J. C./Ito, M./Szentagothai, J.: *The Cerebellum as a Neuronal Machine*, New York 1967.

Estes, W. K.: "The Statistical Approach to Learning Theory", in: S. Koch (ed.), *Psychology. A Study of a Science,* vol. I/2, New York 1959, 380–491.

Fitzhugh, H. S. II: „Some Considerations of Polystable Systems" *IEEE Transactions* 7 (1963), 1–9.

Gill, A.: *Introduction to the Theory of Finite State Machines,* New York 1962.

Günther, G.: "Time, Timeless Logic and Self-Referential Systems", in: R. Fischer (ed.), *Interdisciplinary Perspectives of Time,* New York 1967, 306–406.

Heisenberg, W.: „Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik", *Z. Physik* 43 (1927), 172–198.

Hoagland, H.: "Consciousness and the Chemistry of Time", in: H. A. Abramson (ed.), *Problems of Consciousness (Transactions of the First Conference)*, New York 1951, 164–198.

Hoagland, H.: "(A Remark)", in: H. A. Abramson (ed.), *Problems of Consciousness (Transactions of the Fourth Conference)*, New York 1954, 106–109.

Hubel, D. H./Wiesel, T. N.: "Receptive Fields of Single Neurons in the Cat's Striate Cortex", *Journal of Physiology* 148 (1959), 547–595.

Hubel, D. H./Wiesel, T. N.: "Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex", *Journal of Physiology* 160 (1962), 106–154.

Inselberg, A./von Foerster, H.: *Property Extraction in Linear Networks, Technical Report No. 2,* Electrical Engineering Research Laboratory, Engineering Experiment Station, Urbana/Ill. 1962.

Jean, M.: *Histoire de la Peinture Surrealiste*, Paris 1959.

John, E. R./Shimkochi, M./Bartlett, F.: "Neural Readout from Memory During Generalization", *Science* 164 (1969), 1534–1536.

Katz J. J.: *The Problem of Induction and its Solution,* Chicago 1962.

Konorski, J.: "The Role of Central Factors in Differentiation", in: R. W. Gerard/J. W. Duyff (eds.), *Information Processing in the Nervous System,* vol. 3, Amsterdam 1962, 318–329.

Langer, S.: *Philosophy in a New Key*, New York 1951.

Lettvin, J. Y./Maturana, H. R./McCulloch, W. S./Pitts, W.: "What the Frog's Eye Tells the Frog's Brain", *Proceedings of the Institute of Radio Engineers* 47 (1959), 1940–1951.

Löfgren, L.: "Kinematic and Tesselation Models of Self-Repair", in: E. E. Bernard/M. R. Kare (eds.), *Biological Prototypes and Synthetic Systems*, New York 1962, 342–369.

Löfgren, L.: "Recognition of Order and Evolutionary Systems", in: J. Tou (ed.), *Computer and Information Sciences II,* New York 1967, 165–175.

Löfgren, L.: "An Axiomatic Explanation of Complete Self-Reproduction", *Bulletin of Mathematical Biophysics* 30 (1968), 3, 415–425.

Logan, F. A.: "The Hull-Spence Approach", in: S. Koch (ed.), *Psychology. A Study of a Science,* vol. I/2, New York 1959, 293–358.

Mathiot, M.: "The Semantic and Cognitive Domains of Language", in: P. L. Garvin (ed.): *Cognition: A Multiple View,* New York 1970, 249–277.

Maturana, H. R.: "Neurophysiology of Cognition", in: P. L. Garvin (ed.): *Cognition: A Multiple View*, New York 1970a, 3–23.

Maturana, H. R.: *Biology of Cognition, BCL Report No. 9.0,* Biological Computer Laboratory, University of Illinois, Urbana/Ill. 1970b. (Deutsch „Biologie der Kognition" in: H. R. Maturana, *Erkennen. Die Organisation und Verkörperung von Wirklichkeit,* Braunschweig/Wiesbaden 1982, 32–80.)

Maturana, H. R./Frenk, S.: "Directional Movement and Horizontal Edge Detectors in the Pigeon Retina", *Science* 142 (1963), 977–978.

Maturana, H. R./Uribe, G./Frenk, S.: *A Biological Theory of Relativistic Colour Coding in the Primate Retina*, Archivos de Biologia y Medicina Experimentales, Suppl. No. 1, Santiago 1968. (Deutsch „Eine biologische Theorie der relativistischen Farbkodierung in der Primatenretina" in: H. R. Maturana, *Erkennen. Die Organisation und Verkörperung von Wirklichkeit*, Braunschweig/Wiesbaden 1982, 88–137.)

Maturana, H. R./Varela, F.: *Autopoiesis*, Santiago 1972. (Deutsch „Autopoietische Systeme. Eine Bestimmung der lebendigen Organisation" in: H. R. Maturana, *Erkennen. Die Organisation und Verkörperung von Wirklichkeit*, Braunschweig/Wiesbaden, 170–235.)

McCulloch, W. S.: "A Heterarchy of Values Determined by the Topology of Nervous Nets", *Bulletin of Mathematical Biophysics* 7 (1945), 89–93.

McCulloch, W. S./Pitts, W. H.: "A Logical Calculus of the Ideas Immanent in Nervous Activity", *Bulletin of Mathematical Biophysics* 5 (1943), 115–133.

Miller, G. A.: "Psycholinguistic Approaches to the Study of Communication", in: D. L. Arm (ed.), *Journeys in Science*, Albuquerque: Univ. Mexico 1967, 22–73.

Mountcastle, V. B.: *Abstract in Symposium on Information Processing in the Nervous System,* 22nd International Congress of Physiological Science, Amsterdam 1963.

Naeser, M. A./Lilly, J. C.: "The Repeating Word Effect: Phonetic Analysis of Reported Alternates", *Journal of Speech and Hearing Research* 14 (1971), 32.

National Academy of Science Publication No. 1707, Washington/D.C.: National Academy of Science Printing and Publishing Office 1969, 336 pp.

Pask, G. A.: "The Natural History of Networks", in: M. C. Yovits/S. Cameron (eds.), *Self-Organizing Systems*, London 1960, 232–263.

Pask, G. A.: "A Proposed Evolutionary Model", in: H. von Foerster/G. W. Zopf jr. (eds.), *Principles of Self-Organization,* New York 1962, 229–254.

Pask, G. A.: "A Cybernetic Model for Some Types of Learning and Mentation", in: H. L. Oesterreicher/D. R. Moore (eds.), *Cybernetic Problems in Bionics*, New York 1968, 531–586.

Pattee, H. H.: "On the Origin of Macro-Molecular Sequences", *Biophysical Journal* 1 (1961), 683–710.

Piaget, J.: *The Construction of Reality in the Child*, New York 1954.

Piaget, J./Inhelder, B.: *The Child's Conception of Space*, New York 1956.

Piaget, T.: *L'Equilibration des Structures Cognitives*, Paris 1975.

Pitts, W./McCulloch, W. S.: "How We Know Universals; The Perception of Auditory and Visual Forms", *Bulletin of Mathematical Biophysics* 9 (1947), 127–143.

Planck, M.: „Über eine Verbesserung des Wien'schen Strahlungsgesetzes", in: *Sitzber. Preuss. Ac. Wiss.* Oktober 1900a.

Planck, M.: „Über die Theorie des Energieverteilungsgesetzes normaler Spectren", in: *Sitzber. Preuss. Ac. Wiss.* Dezember 1900b.

Quastler, H.: "A Primer on Information Theory", in: H. P. Yockey/R. L. Platzman/H. Quastler (eds.), *Symposium on Information Theory in Biology*, New York 1958, 3–49.

Schrödinger, E.: *What is Life?*, New York 1947.

Selfridge, O. G.: "The Organization of Organization", in: M. C. Yovits/G. T. Jacoby/G. D. Goldstein (eds.), *Self-Organizing Systems*, New York 1962, 1–8.

Shannon, C. E.: "Prediction and Entropy in Printed English", *The Bell Syst. Tech. J.* 30 (1951), 50–64.

Shannon, C. E./Weaver, W.: *The Mathematical Theory of Communication*, Urbana 1949.

Sherrington, C. S.: *Integrative Action of the Nervous System,* New Haven 1906.

Sholl, D. A.: *The Organization of the Cerebral Cortex*, London 1956.

Skinner, B. F.: "A Case History in Scientific Method", in: S. Koch (ed.), *Psychology. A Study of a Science*, vol. I/2, New York 1959, 359–379.

Skinner, B. F.: *Beyond Freedom and Dignity*, New York 1971.

Teuber, H. L.: „Neuere Betrachtungen über Sehstrahlung und Sehrinde", in: R. Jung/H. Kornhuber (eds.): *Das visuelle System*, Berlin 1961, 256–274.

TIME Magazine 5 January 1970, "The Middle Americans".

Ungar, G.: "Chemical Transfer of Learning", in: S. Bogoch (ed.), *The Future of the Brain Sciences*, New York 1969, 373–374.

von Foerster, H.: *Das Gedächtnis. Eine quantenmechanische Untersuchung*, Wien 1948.

von Foerster, H. (ed.): *Cybernetics: Circular Causal and Feedback Mechanisms in Biological and Social Systems,* Transactions of the Sixth Conference, Josiah Macy Jr. Foundation, New York 1949.

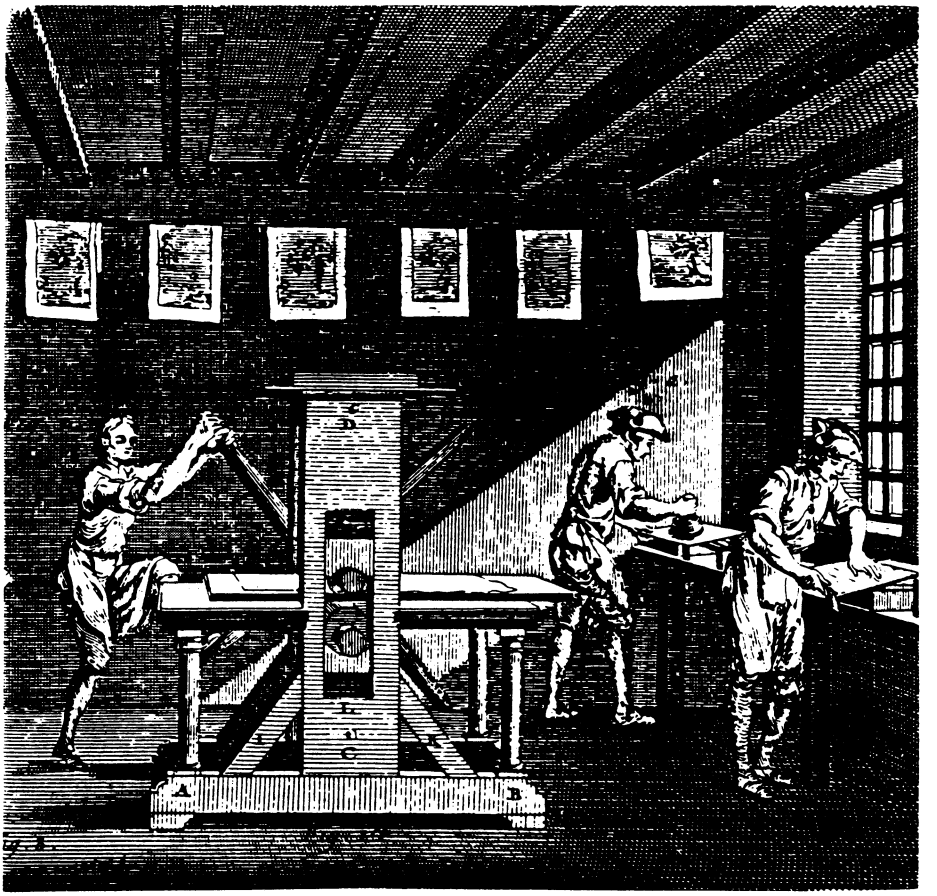
von Foerster, H.: "Quantum Mechanical Theory of Memory", *ibid.* 112–145.
von Foerster, H.: "On Self-Organizing Systems and their Environment", in: M. C. Yovits/S. Cameron (eds.): *Self-Organizing Systems*, New York 1960, 31–50. (Deutsch „Über selbstorganisierende Systeme und ihre Umwelten" *in diesem Band.*)
von Foerster, H.: "Molecular Bionics", in: H. L. Oesterreicher (ed.), *Information Processing by Living Organisms and Machines*, Dayton 1964, 161–190.
von Foerster, H.: "Memory without Record", in: D. P. Kimble (ed.): *The Anatomy of Memory*, Palo Alto 1965, 388–433. (Deutsch „Gedächtnis ohne Aufzeichnung" *in diesem Band.*)
von Foerster, H.: "From Stimulus to Symbol", in: G. Kepes (ed.): *Sign, Image, Symbol*, New York 1966, 42–61.
von Foerster, H.: "Biological Principles of Information Storage and Retrieval", in: A. Kent (ed.), *Electronic Handling of Information: Testing and Evaluation*, London 1967a, 123–147.
von Foerster, H.: "Computation in Neural Nets", *Currents Mod. Biol.* 1 (1967b), 47–93.
von Foerster, H.: "What is Memory that it may have Hindsight and Foresight as well?", in: S. Bogoch (ed.), *The Future of the Brain Sciences*, New York 1969, 19–64.
von Foerster, H.: "Thoughts and Notes on Cognition", in: P. Garvin (ed.): *Cognition: A Multiple View*, New York 1970a, 25–48. (Deutsch „Gedanken und Notizen über Kognition" *in diesem Band.*)
von Foerster, H.: "Molecular Ethology: An Immodest Proposal for Semantic Clarification", in: G. Ungar (ed.), *Molecular Mechanisms in Memory and Learning*, New York 1970b, 213–248. (Deutsch „Molukular-Ethologie: ein unbescheidener Versuch semantischer Klärung" *in diesem Band.*)
von Foerster, H./Inselberg, A./Weston, P.: "Memory and Inductive Inference", in: H. L. Oesterreicher/D. R. Moore (eds.), *Cybernetic Problems in Bionics*, New York 1968, 31–68.
von Foerster, H./Mead, M./Teuber, H. L. (eds.): *Cybernetics: Circular Causal and Feedback Mechanisms in Biological and Social Systems*, Transactions of the Seventh Conference, Josiah Macy jr. Foundation, New York 1950.
von Foerster, H./Mead, M./Teuber, H. L. (eds.): *Cybernetics: Circular Causal and Feedback Mechanisms in Biological and Social Systems*, Transactions of the Eighth Conference, Josiah Macy jr. Foundation, New York 1951.
von Foerster, H./Mead, M./Teuber, H. L. (eds.): *Cybernetics: Circular Causal and Feedback Mechanisms in Biological and Social Systems*, Transactions of the Ninth Conference, Josiah Macy jr. Foundation, New York 1953.
von Foerster, H./Mead, M./Teuber, H. L. (eds.): *Cybernetics: Circular Causal and Feedback Mechanisms in Biological and Social Systems*, Transactions of the Tenth Conference, Josiah Macy jr. Foundation, New York 1955.
von Neumann, J.: *The Theory of Self-Reproducing Automata*, Urbana/Ill. 1966.
Walker, C.: *A Study of a Family of Complex Systems.* An Approach to the Investigation of an Organism's Behavior, Technical Report 5, Biological Computer Laboratory, Urbana/Ill. 1965.
Werner, G.: "The Topology of the Body Representation in the Somatic Afferent Pathways", in: *The Neurosciences II* (ed. F. O. Schmitt), New York 1970, 87–102.
Weston, P.: *Noun Chain Trees*, Unpubl. Manuscript 1964.
Weston, P.: "To Uncover; To Deduce; To Conclude", *Comp. Stud. in the Humanities and Verbal Behavior* vol. 3(2), 1970a, 77–89.
Weston, P.: *CYLINDERS. A Relational Data Structure.* Techn. Rep. No. 18, AFOSR 70–1865, Biological Computer Laboratory, Electrical Engineering Research Laboratory, Engineering Experiment Station, University of Illinois, Urbana 1970b, 78 pp.
Weston, P.: "An End of Search, A Means of Understanding: A Preface to the Antidocument", in: *Accomplishment Summary*, BCL Rep. No. 71.2, Biological Computer Laboratory, Department of Electrical Engineering, University of Illinois, Urbana 1971.

Wittgenstein, L.: *Tractatus Logico Philosophicus*, New York 1961.

Witz, K.: *Models of Sensory-Motor Schemes in Infants*, Research Report, Department of Mathematics, University of Illinois, Urbana/Ill. 1972.

Witz, K./Easley, J.: "Cognitive Deep Structure and Science Education", *Final Report. Analysis of Cognitive Behavior in Children*, Curriculum Laboratory, University of Illinois, Urbana/Ill. 1972.

Worden, F. G.: "EEG Studies and Conditional Reflexes in Man", in: M. A. B. Brazier (ed.), *The Central Nervous System and Behavior*, New York 1959, 270–291.


D
E
C
A
B

Veröffentlichungen

Heinz von Foerster

„Über das Leistungsproblem beim Klystron", *Ber. Lilienthal Ges.* Luftfahrtforschung 155 (1943), 1–5.

Das Gedächtnis: Eine quantenmechanische Untersuchung, Wien 1948.

(Ed.): *Cybernetics: Transactions of the Sixth Conference*, Josiah Macy jr. Foundation, New York 1949.

"Quantum Mechanical Theory of Memory", *ibid.* 112–145.

(Ed. mit M. Mead und H. L. Teuber): *Cybernetics: Transactions of the Seventh Conference*, Josiah Macy jr. Foundation, New York 1950.

(Ed. mit M. Mead und H. L. Teuber): *Cybernetics: Transactions of the Eighth Conference*, Josiah Macy jr. Foundation, New York 1951.

Mit M. L. Babcock und D. F. Holshouser: "Diode Characteristics of a Hollow Cathode", *Phys. Rev.* 91 (1953), 755.

(Ed. mit M. Mead und H. L. Teuber): *Cybernetics: Transactions of the Ninth Conference*, Josiah Macy jr. Foundation, New York 1953.

Mit E. W. Ernst: "Electron Bunches of Short Time Duration", *J. of Appl. Phys.* 25 (1954), 674.

Mit L. R. Bloom: "Ultra-High Frequency Beam Analyzer", *Rev. of Sci. Instr.* 25 (1954), 640–653.

„Experiment in Popularization", *Nature* 174 (1954), 4424.

(Ed. mit M. Mead und H. L. Teuber): *Cybernetics: Transactions of the Tenth Conference*, Josiah Macy jr. Foundation, New York 1955.

Mit O. T. Purl: "Velocity Spectrography of Electron Dynamics in the Traveling Field", *Journ. of Appl. Phys.* 26 (1955), 351–353.

Mit E. W. Ernst: "Time Dispersion on Secondary Electron Emission", *Journ. of Appl. Phys.* 26 (1955), 781–782.

Mit M. Weinstein: "Space Charge Effects in Dense, Velocity Modulated Electron Beams", *Journ. of Appl. Phys.* 27 (1956), 344–346.

Mit E. W. Ernst, O. T. Purl und M. Weinstein: "Oscillographie Analyse d'un Faisceau Hyperfréquences", *Le Vide* 70 (1957), 341–351.

"Basic Concepts of Homeostasis", *Homeostatic Mechanisms*, New York 1958, 216–242.

Mit G. Brecher und E. P. Cronkite: „Produktion, Ausreifung und Lebensdauer der Leukozyten", in: H. Braunsteiner (ed.), *Physiologie und Physiopathologie der weißen Blutzellen*, Stuttgart 1959, 188–214.

"Some Remarks on Changing Populations", in: F. Stohlman jr. (ed.): *The Kinetics of Cellular Proliferation*, New York 1959, 382–407.

"On Self-Organizing Systems and Their Environments", in M. C. Yovits/S. Cameron (eds.): *Self-Organizing Systems*, London 1960, 31–51. (Russisch: "O samoorganizuyushchiesja sistemach i ich okrooshenii", in: *Samoorganizuyushchiesju Sistemi*, Moskau 1964, 113–139. Spanisch "Sobre sistemas autoorganizados y sus contornos", in: *Epistemologia de la Comunicacion* ed. J. A. Bofil, Valencia 1976, 187–214.)

Mit P. M. Mora und L. W. Amiot: "Doomsday: Friday, November 13, AD 2026", *Science* 132 (1960), 1291–1295.

"Bionics", in: J. Steele (ed.), *Bionics Symposium*, Wright Air Development Division, Technical Report 60–600, 1960, 1–4.
"Some Aspects in the Design of Biological Computers", *Second International Congress on Cybernetics*, Namur 1960, 241–255.
Mit G. Pask: "A Predictive Model for Self-Organizing Systems", Teil I: *Cybernetica* 3(1961), 258–300; Teil II: *Cybernetica* 4 (1961), 20–55.
Mit P. M. Mora und L. W. Amiot: "Doomsday", *Science* 133 (1961), 936–946.
Mit D. F. Holshouser und G. L. Clark: "Microwave Modulation of Light Using the Kerr Effect", *Journ. Opt. Soc. Amer.* 51 (1961), 1360–1365.
Mit P. M. Mora und L. W. Amiot: "Population Density and Growth", *Science* 133 (1961), 1931–1937.
Mit G. Brecher und E. P. Cronkite: "Production, Differentiation and Lifespan of Leukocytes", in: H. Braunsteiner (ed.): *The Physiology and Pathology of Leukocytes*, New York 1962, 170–195.
(Ed. mit G. W. Zopf jr.): *Principles of Self-Organization. The Illinois Symposium on Theory and Technology of Self-Organizing Systems*, London 1962.
"Communication Amongst Automata", *Amer. Journ. Psychiatry* 118 (1962), 865–871.
Mit P. M. Mora und L. W. Amiot: "'Projections' versus 'Forecasts' in Human Population Studies", *Sience* 136 (1962), 173–174.
"Biological Ideas for the Engineer", *The New Scientist* 15 (1962), 173–174.
"Bio-Logic", in: E. E. Bernard/M. A. Kare (eds.), *Biological Prototypes and Synthetic Systems*, New York 1962, 1–12. (Russisch "Bio-Logika", in: *Problemi Bioniki*, Moskau 1965, 9–23.)
"Circuitry of Clues of Platonic Ideation", in: C. A. Muses (ed.), *Aspects of the Theory of Artificial Intelligence*, New York 1962, 43–82.
"Perception of Form in Biological and Man-Made Systems", in: E. J. Zagorski (ed.), *Trans. I.D.E.A. Symp.*, Urbana/Ill. 1962, 10–37.
Mit W. R. Ashby und C. C. Walker: "Instability of Pulse Activity in a Net with Threshold", *Nature* 196 (1962), 561–562.
"Bionics", in: *McGraw-Hill Yearbook Science and Technology*, New York 1963, 148–151.
"Logical Structure of Environment and its Internal Representation", in: R. E. Eckerstrom (ed.): *Transactions of the International Design Conference Aspen*, Zeeland/Mich. 1963, 27–38.
Mit W. R. Ashby und C. C. Walker: "The Essential Instability of Systems with Threshold, and Some Possible Applications to Psychiatry", in: N. Wiener/J. P. Schade (eds.), *Nerve, Brain and Memory Models*, Amsterdam 1963, 136–243.
"Molecular Bionics", in: H. L. Oesterreicher (ed.), *Information Processing by Living Organisms and Machines*, Aerospace Medical Division, Dayton 1964, 161–190.
Mit W. R. Ashby: "Biological Computers", in: K. E. Schaefer (ed.), *Bioastronautics*, New York 1964, 333–360.
"Form: Perception, Representation and Symbolization", in: N. Perman (ed.), *Form and Meaning*, Chicago 1964, 21–54.
"Structural Models of Functional Interactions", in: R. W. Gerard/J. W. Duyff (eds.), *Information Processing in the Nervous System*, Amsterdam 1964, 370–383.
"Physics and Anthropology", *Current Anthropology* 5 (1964), 330–331.
"Memory without Record", in: D. P. Kimble (ed.), *The Anatomy of Memory*, Palo Alto 1965, 288–433.
"Bionics Principles", in: R. A. Willaume (ed.), *Bionics*, Paris 1965, 1–12.
"From Stimulus to Symbol", in: G. Kepes (ed.), *Sign, Image, Symbol*, New York 1966, 42–61.
"Computation in Neural Nets", *Currents Mod. Biol.* 1 (1967), 47–93.
"Time and Memory", in: R. Fischer (ed.), *Interdisciplinary Perspectives of Time*, New York 1967, 866–873.

Mit G. Günther: "The Logical Structure of Evolution and Emanation", in: R. Fischer (ed.), *Interdisciplinary Perspectives of Time*, New York 1967, 874–891.

"Biological Principles of Information Storage and Retrieval", in: A. Kent (ed.), *Electronic Handling of Information: Testing and Evaluation*, London 1967, 123–147.

Mit A. Inselberg und P. Weston: "Memory and Inductive Inference", in: H. Oesterreicher/D. Moore (eds.), *Cybernetic Problems in Bionics*, New York 1968, 31–68.

(Ed. mit J. White, L. Peterson und J. Russell): *Purposive Systems*, Proceedings of the 1st Annual Symposium of the American Society for Cybernetics, New York 1968.

(Ed. mit J. W. Beauchamp): *Music by Computers*, New York 1969.

"Sounds and Music", *ibid.*, 3–10.

"What is Memory that it may have Hindsight and Foresight as well?" in: S. Gogoch (ed.), *The Future of the Brain Sciences*, New York 1969, 19–64.

"Laws of Form" (Rezension von G. S. Brown, *Laws of Form*), *Whole Earth Catalog* 14, Palo Alto 1969.

"Molecular Ethology: An Immodest Proposal for Semantic Clarification", in: G. Ungar (ed.), *Molecular Mechanisms in Memory and Learning*, New York 1970, 213–248.

Mit A. Inselberg: "A Mathematical Model of the Basilar Membrane", *Mathematical Biosciences* 7 (1970), 341–363.

"Thoughts and Notes on Cognition", in: P. Garvin (ed.), *Cognition. A Multiple View*, New York 1970, 25–48.

"Bionics, Critique and Outlook", in: H. E. von Gierke/W. D. Keidel/H. L. Oesterreicher (eds.), *Principles and Practice of Bionics*, Slough 1970, 467–473.

"Embodiments of Mind" (Rezension von W. S. McCulloch, *Embodiments of Mind*), *Computer Studies in the Humanities and Verbal Behavior* 3 (1970), 2, 111–112.

Mit L. Peterson: "Cybernetics of Taxation: The Optimization of Economic Participation", *Journal of Cybernetics* 1 (1970), 2, 5–22.

"Obituary for Warren S. McCulloch", *ASC Newsletter* 3 (1970), 1.

"Preface", in: S. Chermayeff/A. Tzonis (eds.), *Shape of Community*, Baltimore 1971, XVII–XXI.

(Ed.): *Interpersonal Relational Networks*, CIDOC Cuaderno No. 1014, Centro Intercultural de Documentacion, Cuernavaca 1971.

"Technology: What will it Mean to Librarians?", *Illinois Libraries* 53 (1971), 9, 785–803.

"Computing in the Semantic Domain", *Annals of the New York Academy of Science* 184 (1971), 239–241.

"Responsibilities of Competence", *Journal of Cybernetics* 2 (1972), 2, 1–6.

"Perception of the Future and the Future of Perception", *Instructional Science* 1 (1972), 1, 31–43. (Spanisch "La Percepción de Futuro y el Futuro de Percepción", *Communicación*, Barcelona 1975.)

Notes on an Epistemology for Living Things, BCL Report No. 9.3, Biological Computer Laboratory, University of Illinois, Urbana/Ill. 1972. (Französisch "Notes pour une épistémologie des objets vivants" in: E. Morin/M. Piatelli-Palmarini (eds.), *L'unité de l'homme*, Paris 1974, 401–417.)

Mit P. E. Weston: "Artificial Intelligence and Machines that Understand", in: H. Eyring/C. J. Christensen/H. S. Johnston (eds.), *Annual Review of Physical Chemistry*, Palo Alto 1973, 353–378.

"On Constructing a Reality", in: F. E. Preiser (ed.), *Environmental Design Research*, vol. 2, Stroudberg 1973, 35–46. (Spanisch "Construir la realidad", in: *Infancia y Aprendizaje* 1/1, Madrid 1978, 79–92. Deutsch „Das Konstruieren einer Wirklichkeit", in: *Die erfundene Wirklichkeit* hrsg. von P. Watzlawick, München 1981, 39–60.)

Mit P. Arnold, B. Aton, D. Rosenfeld und K. Saxena: "Diversity: H-A Measure Complementing Uncertainty H", *Systema* 2 (1974).

Giving with a Purpose: The Cybernetics of Philanthropy, Occasional Paper No. 5, Washington 1974.

„Kybernetik einer Erkenntnistheorie", in: W. D. Keidel/W. Händler/M. Spreng (eds.), *Kybernetik und Bionik*, München 1974, 27–46.

"Epilogue to Afterwords", *After Brookman: A Symposium*, ABYSS, 4 (1974), 68–69.

Mit R. Howe: "Cybernetics at Illinois", Teil I: *Forum* 6 (1974), 3, 15–17; Teil II: *Forum* 6 (1974), 4, 22–28.

"Culture and Biological Man" (Rezension von *Culture and Biological Man* von Elliot D. Chapple), *Current Anthropology* vol. 15 (1974), 61.

"Comunicación, autonomía y libertad" (Entrovista con H. v. Foerster), *Comunicaión* no. 14, Madrid 1974, 33–37.

Mit R. Howe: "Introductory Comments to Francisco Varela's 'Calculus for Self-Reference'", *Int. J. General Systems* 2 (1975), 1–3.

"Two Cybernetics Frontiers" (Rezension von S. Brand, *Two Cybernetics Frontiers*), *The Co-Evolutionary Quarterly* 2 (Summer 1975), 143.

"Oops; Gaia's Cybernetics Badly Expressed", *The Co-Evolution Quarterly* 2 (1975), 51.

"Objects: Tokens for (Eigen)Behaviors", *ASC Cybernetics Forum* 8/3&4, 1976, 91–96. (Französisch: "Formalisation de certains aspects des l'équilibration des structures cognitives", in: *Epistémologie génétique et équilibration* ed. B. Inhelder/R. Garcia/J. Voneche, Neuchatel 1977, 76–89.)

"On 'Where Do We Go From Here?'", in: G. Bugliarello (ed.), *Proceedings of the International Symposium on the History and Philosophy of Technology*, Urbana/Ill. 1978.

"Cybernetics of Cybernetics", in: K. Krippendorf (ed.), *Communication and Control in Society*, New York 1978, 1–4.

"Minicomputer – verbindende Elemente", *Chip*, Januar 1980, 8.

"Epistemology of Communication", in: *The Myths of Information. Technology and Post-Industrial Culture* ed. K. Woodward, Madison 1980, 18–27.

"Morality Play", *The Sciences* 21/8, 1981, 24–25.

"Gregory Bateson", *The Esalen Catalogue* 20/1, 1981, 10.

"On Cybernetics of Cybernetics and Social Theory", in: *Self-Organizing Systems* eds. G. Roth/H. Schwegler, Frankfurt/M. 1981, 102–105.

"Foreword" to *Rigor and Imagination* eds. C. Wilder-Mott/J. H. Weakland, New York 1981, vii-xi.

"Understanding Understanding: An Epistemology of Second-Order Concepts", in: *Aprendizagem/Desenvolvimento* 1/3, 1981, 83–85.

Observing Systems, Seaside/Cal. 1982.

"A Constructivist Epistemology", *Cahiers de la Fondation Archives Jean Piaget* no. 3, Genf 1982, 191–213.

The Needs of Perception for the Perception of Needs, American Institute of Architects, Atlanta 1975.

"Second Order Concepts: An Elliptical Parable of Circular Causality", *General Systems Bulletin* 7 (1977), 2, 7–11.

"The Curious Behavior of Complex Systems: Lessons from Biology", in: H. A. Linstone/W. H. C. Simmonds (eds.), *Futures Research*, Reading 1977, 104–113.

"Culture and Biological Man" (Rezension von E. D. Chapple, *Culture and Biological Man*), *Current Anthropology* 15 (1974), 1, 61.

"Sobre Sistemas Autoorganizados y sus Contornos", in: J. A. Bofil (ed.): *Epistemología de la Communicación*, Valencia 1976, 187–214.

* * * * *

(Anm. d. Übs.: Die Arbeiten H. von Foersters sowie aller Mitarbeiter des von ihm aufgebauten und geleiteten "Biological Computer Laboratory (BCL)" der University of Illinois, ca. 14.000 Seiten, sind auf 146 Microfiches (in einer Mappe zusammengefaßt) verfügbar: K. L. Wilson (ed.): *The Collected Works of the Biological Computer Laboratory*, Department of Electrical Engineering, University of Illinois, Urbana/Ill., USA, 1976.

Zu bestellen bei: Illinois Blueprint Corp., Micrographics Dept., 821 Bond, Peoria, Illinois 61603, U.S.A. Preis $ 59.00 zuzgl. $ 5.00 für Porto nach Übersee.)

Namenverzeichnis

Aristophanes 5
Aristoteles 9, 12, 62, 85
Arnold, P. 73, 75
Ashby, W. R. 112, 177, 185 f
Babcock, M. L. 158
Bielby, W. 53
Boltzmann, L. 119
Bower, T. G. R. 72
Brillouin, L. 139
Brown, G. S. 25, 47
H. Brün 3, 13
Buber, M. 79
Bullock, T. H. 198
Dali, S. 75, 96
Dancoff, S. M. 154
Defares, J. G. 139
Descartes, R. 36
Easley, J. 83
Ebbinghaus, H. 200
Eccles, J. 135, 192
Einstein, A. 81
Estes, W. K. 187, 191
Fitzhugh, H. S. 185
Frankl, V. 79
Frenk, S. 158
Gabor, D. 23
Gill, A. 177
Gödel, K. 21
Günther, G. 21, 112, 176
Heisenberg, W. 123
Hoagland, H. 200
Hubel, D. H. 158, 160
Inhelder, B. 72
Inselberg, A. 48, 158
Jaspers, K. 79
John, E. R. 186
Katz, J. J. 47, 66
Konorski, J. 92, 111, 176
Kopernikus, N. 119
Langer, S. K. 92, 110
Lettvin, J. Y. 158, 192
Lilly, J. C. 27
Löfgren, L. 19, 21, 47, 69, 176, 196
Logan, F. A. 187
Lorenz, K. 96
Mathiot, M. 97
Maturana, H. R. 4, 31, 46, 79, 85, 90, 98, 100, 112, 158, 175 f., 192
McCulloch, W. S. 175, 198, 208, 215
Mead, M. 66
Miller, G. 6
Mountcastle, V. B. 158
Naeser, M. A. 27
Neumann, J. von 177, 196
Pask, G. 117 f., 196
Piaget, J. 72, 83, 207
Pitts, W. 175, 198, 208
Planck, M. 44
Quastler, H. 111, 139, 144, 154
Schrödinger, E. 124 f.
Shannon, C. E. 120, 139, 147
Sherrington, C. S. 199
Sholl, D. A. 35
Seligman, K. 95
Skinner, B. F. 187
Sneddon, I. N. 139
Teuber, H. L. 27, 66
Varela, F. 85
Walker, C. 185
Wallis, 88
Weaver, W. 139
Weston, P. 48, 50, 52, 61, 174
Wiener, N. 18, 66
Wiesel, T. N. 158
Wittgenstein, L. 92, 110, 119
Witz, K. 72, 83
Woods, L. 112
Worden, F. G. 27

Sachwortverzeichnis

Abstraktion 84, 87, 160 f.
–, Berechnung von – 160
Äquivalenzrelation 84, 86
ästhetischer Imperativ 41
äußerer Dämon 123
Aktionsfunktion 159 ff., 163, 165 f.
Algorithmus, deduktiver 175
–, induktiver 175
Alternanten 27
Anthropomorphisierung 95
Anthropomorphismus 96 f.
Antriebsfunktion 178, 180 ff., 183
Aufzeichnung 133, 138, 175, 200
autologisches Koordinatensystem 70 f.
autonome Elemente 196
Autonomie 13, 40, 46, 79
Autonomie-Eigenschaften 196
Autopoiese 85
Axiome der zweiwertigen Logik 91, 109
Axon 33

Bahnungseffekt 35
Bedeutung 40
bedingter Reflex 92, 110
Beobachter 70 f., 77, 81 f., 85, 90 ff., 111, 119, 175, 197, 207 f., 212
Beobachtung 91
Beschreibung 20, 31, 45 f., 68, 81, 83, 85, 89 ff.
– der Umwelt 109
–, Information einer – 93
– von Beschreibungen 110
– von Tatsachen 204
Bewußtsein 92, 110
Bibliothek 45 f., 58
– der Zukunft 44
Bibliothekare 43 f.
Bildungssystem (Erziehung) 13
Biophysik 197
Bit 144
blinder Fleck 26
Buch 45, 58, 62

Computer 96

Dämonen 124, 138
–, äußere 123
–, innere 122
Dämon, Maxwellscher 18 ff., 123, 146
Deduktion 85, 93
deduktiver Algorithmus 175
Definitionsstruktur, hierarchische 174 f.
–, heterarchische 174
Dendriten 33
deterministische Maschine 12, 177
Dysgnosie 10 f.

Eigenfunktion 193
Eigen-Koordinatensystem 102 ff., 108
(Eigen-)Verhalten 207, 210, 213
Eigenwert 66, 72, 79, 211 f., 214 ff.
Eigen-Werte 210
Einheiten der Dokumentation 58
Empfindlichkeit 77
Entropie 19, 116, 119, 121 ff., 126 ff., 146, 190 f., 202
–, negative 125
–, relative 120
Entropieverzögerung 19, 21
Entschlüsselung des genetischen Kodes 198
epistemische Homöostase, Postulat der 79
Epistemologie 65, 81 f., 207
Ereignis 84, 87, 90
Er-Kennen 67, 69
Erkenntnis 4, 65, 69
Erkenntnistheorie 65, 67
Er-Rechnen 67 ff.
(Er-)Rechner 138
– induktiver Schlüsse 153
Errechnung 30 f., 36 ff., 70, 88, 161, 173, 193
Erziehung (Bildung) 9, 21 f.
Ethik 79, 213
ethischer Imperativ 41
Ethologie 47, 173
–, molekulare 173
Evolution 31

Formalismen, geschlossene 47
Formberechnung 37
Frage-Antwort-System 52, 57
Funktionale (Relationen zw. Funktionen) 192
Funktionen, logische 198
Futurologen 10

Gedächtnis 48, 84, 87, 97 f., 133, 138, 152 f., 160, 168, 175 f., 181, 187, 191 ff., 197, 203 f.
–, kognitives 52, 55 ff.
Gegenstände (Objekte) 207
Gehirn 21, 33, 48, 134, 153 f.
genetische Information 157
genetischer Kode, Entschlüsselung 198
genetisches Programm 161 f., 169 f.
–, Informationsgehalt 154 f.
geschlossene Systeme 65
Geschlossenheit (Zirkularität) 39, 211, 213
Grenze unseres Systems 119 f.
Großinquisitor 14

2. Hauptsatz der Thermodynamik 19, 115 f., 146
Heinz von Foersters Theorem Nr. 1 17
– Theorem Nr. 2 17
– Theorem Nr. 3 21
Heterarchien 215
heterarchische Definitionsstruktur 174
heterologisches Koordinatensystem 71
Heteronomie 40, 79
hierarchische Struktur 175
Homöostase, epistemische 79

Ich-Bewußtsein 90
Identität 41
illegitime Fragen 13, 22
Indexiersprache 57, 59
Induktion 85, 93
induktiver Algorithmus 175
induktive Schlüsse 138
Inferenzmaschine, induktive 200
Information 4 f., 47, 66, 85, 92, 97, 99, 110 ff., 134, 169
– einer Beschreibung 93
–, genetische 157
Informationsbetrag 111, 146, 153
Informationsbetrag H 92
Informationsgehalt des genetischen Programms 154 f.
Informationstheorie 139
inhibitorischer Effekt 35
innerer Dämon 122
Inputzustände 177 ff.
Intelligenz, künstliche 97
interagierende Maschinen 183
Interaktion des Organismus 183
– mit anderen Systemen 194
Invarianz 83, 86 f., 160

Kalkül rekursiver Ausdrücke 85
Kaskade 155 ff.
Kausalkreis, geschlossener 66
–, Schließung des 79
Kode, genetischer 198
Kognition 30 f., 46 f., 55, 95, 100, 211, 213
Kognition als selbstreferentieller Prozeß 196
–, Problem der 99
kognitive Diversität, Prinzip der 212
kognitive Homöostase 39
kognitive Kontinuität, Prinzip der – 212
kognitive Prozesse 65
kognitiver Rechenprozeß 211
kognitives Gedächtnis 52, 55, 57
Kommunikation 61, 85, 90
–, Theorie 90
–, Technologie 90
Kommunikationsviskositätskonstante 61
konnotative Sprache 85
Koordinatensystem, autologisches 70 f.
–, heterologisches 71
–, subjektives 72 f.
kreiskausale Phänomene 65
kreiskausal geschlossene Mechanismen 66
künstliche Intelligenz 97
Kybernetik 18 f., 21, 23, 65 ff.

legitime Frage 13, 22
Lernen, 98, 136 f., 175 f., 191 f., 203
Lernexperimente 183, 186
Lerntheorie, statistische 187 f., 190
logische Funktionen 198
logische Partikel 92, 110

Maschine, Begriff der – 197
Maschinen, deterministische 12, 177, 187, 191 f.
–, interagierende 183
– mit endlich vielen Funktionen 192 f., 204
– mit endlich vielen Zuständen 177, 184 f., 199
–, nicht-triviale 12, 21, 179 f., 182 f., 187, 194, 196
–, probabilistische 187, 191
–, triviale 179 f., 182 f., 187, 194, 196, 203
Maß für Ordnung 119 f.
mathematische Induktion, Prinzip der – 119
Maxwellscher Dämon 18 ff., 123, 146

Mechanismen, kreiskausal geschlossene 66
Metaprogramm 38, 69, 89
molekulare Ethologie 173
Molekularrechner 200
Molekularspeicher 200
Molekularträger 201
motorische Oberfläche 39
Motorium 47, 66, 72 f.
multistabile Systeme 186

Nahwirkungsgesetz, neurologisches (raum-zeitliches) 77
Name 87
Negation 92, 110
Negentropie 19, 120
Nervennetz 199
Nervensystem 77, 79, 90, 135, 153
Netz, elementares 156 f.
Netzwerk 153, 155, 164
–, adaptives abstrahierendes 161, 167
–, elementares 156, 169
–, Störungen 169
neurologisches (raum-zeitliches) Nahwirkungsgesetz 77
Neuronen 33, 77, 153
–, Anzahl der 154
Neurophysiologie 31 ff.
nicht-lineare Systeme 17
nicht-triviale Maschine 12, 21
nicht-triviale Verhaltensweisen 33
notwendige und hinreichende Ursache, Prinzip der – 9
Notwendigkeit 85, 93

Oberfläche, sensorische 39
–, motorische 39
Objekt(e) 82 ff., 86 f., 89, 93, 208, 212 f.
Objektivität 212
Objektkonstanz 83
offene Systeme 65
Ordnen 21
Ordnung 5 f., 19 f., 30, 139, 152 f., 171
–, Maß der – 146 f.
Ordnung aus Ordnung 124 f.
–, Prinzip der – 128
Ordnung aus Unordnung 124 f.
Ordnung der Gene 124
Ordnung der Umwelt 117 ff.
Ordnung des Systems 121
Ordnung durch Störung 125, 128
–, Prinzip der 128
Organisation des Nervensystems 78
Organismen 19 f.
–, autonome selbst-referentielle 204
–, lernende 138
Organismus 39, 46 f., 77, 84 f., 90 ff., 151 f., 175 f., 200
–, autonome Einheit 100
– und Umwelt 183 f., 192
Outputzustände 177 ff.

pathologische Semantik 99
Prinzip der Erhaltung von Regeln 9, 11
Prinzip der kognitiven Diversität 212
Prinzip der mathematischen Induktion 119
Prinzip der notwendigen und hinreichenden Ursache 9
Prinzip der rekursiven Errechnung 31
Prinzip der Relativität 40 f.
Prinzip der undifferenzierten Kodierung 29
probabilistische Maschinen 187, 191
Problem der Kognition 29 f., 99
Problemlöser 99
Proportionalitätskonstante 159
Programm 38, 89, 173, 192
–, genetisches 154 f., 161 f., 169 f.

4. räumliche Dimension 73, 76
Reaktion eines Organismus 108
Realität 30 f., 39 ff., 67 f., 117 ff.
–, er-rechnete 79
–, externe (objektive) 84
Rechennetz 157
Rechenprozesse 65
–, kognitive 65, 211
–, rekursive 65
Rechnen 19 ff., 30, 33, 83
– im semantischen Bereich 49 f., 52
Rechner 134 f., 137, 153
–, adaptiver 135 f.
Redundanz 120
Reflex, bedingter 92, 110
Regeln 21
Regelung 18 f., 23, 40
Regler 10 f., 79
Rekursion 31, 84, 112, 189, 198, 207, 210
rekursive Ausdrücke 85
– Definition 85
– Organisation 38, 207
– Vorschrift 74
rekursive Errechnung, Prinzip der – 31
rekursives Errechnen 69, 77
rekursiv rechnend 40

Relation 84 ff., 87 ff., 91 f.
– zwischen Organismus und Umwelt 107
Relativität, Prinzip der – 40 f.
Relativitätsprinzip 119
Relator 84 f., 89 f., 91
Repräsentation 83 ff., 86 ff., 90 ff., 192
– der Umwelt 109
– von Repräsentationen 110
repräsentative Einheitskugel 70 f.
– Körperkugel 102 f.
Repräsentoren 84, 89, 91
Retina 36
rezeptives Feld 160
Rückbezüglichkeit 81
(s. auch Selbstreferenz)

Schließung des Kausalkreises 79
Schlüsse, induktive 138
Selbstorganisation 196
selbst-organisierende Systeme 115 ff., 119, 121 ff., 125, 149
selbst-referentiell 196, 215
Selbst-Referenz (Selbstreferenz) 21, 71, 107, 176
Selbst-Regelung
siehe Autonomie
Semantik 49, 176
–, pathologische 99
semantische Falle 47, 175
– Fallstricke 4
– Klärung 173
– Konfusion 4, 98
– Modelle 51
– Rechenprozesse 48, 54
– Verdrehung 99
Sensitivität 35
sensorische Modalität 105, 107
– Oberfläche 39
Sensorium 47, 66, 72 f.
Skotom 26
Solipsismus 40
Speicher 98, 135
Speicherung 49
Sprachen 20
–, konnotative 85
Stimulus 108
Struktur der Umwelt 119
Struktur, hierarchische 175
subjektives Koordinatensystem 72 f.
Synapsen 34, 166
synaptischer Spalt 35
Systeme, geschlossene 65
–, multistabile 186
–, nicht-lineare 17
–, offene 65
System, Grenze unseres – 119 f.
System von Frage und Antwort 52, 57

Technik 43
Technokratie 43
teleologische Theorie der Kommunikation 90
Tesselierung 194 f., 204
Theorem Heinz von Foersters Nr. 1 17
– Nr. 2 17
– Nr. 3 21
Theorie des Beobachters, geschlossene 79
Theorie des Gehirns 21
Theorie des Rechnens 177
Thermodynamik, 2. Hauptsatz der – 19, 115 f., 146
Transferfunktion 163 f., 166, 170
Transferfunktionen eines Neurons 200
triviale Maschine 12, 21
triviale Mechanismen 98
Trivialisierung 12 f., 21
Turingmaschine 21, 69

Umwelt 25, 72, 77, 82, 84 ff., 87, 92 f., 100, 103 ff., 109, 137 f., 152 f., 170
–, Beschreibung der – 109, 112
–, Repräsentation der – 109
Umwelten 20, 115 ff.
undifferenzierte Kodierung 29, 69
Ungewißheit 139
universale Turingmaschine 19 f.
Ursache 9, 66

Verantwortung 40, 43, 79, 99
Verantwortung (des Experten) 17
Verantwortung (durch Wissen) 23
Veränderung 3, 10
–, räumliche 38
– der Erscheinung eines Objekts 82 f., 86 f.
– von Formen 175
Verhalten 203
Verhaltensweisen, nicht-triviale 33
Verstehen 27

Wachstum 151 f.
Wahrnehmen 6, 27, 30, 98, 175
Wahrnehmung der Zukunft 3
Wahrnehmung (eines Organismus) 47
Wirklichkeit 25, 88
Wirkursache 9
Wissenschaftlergemeinschaft 61
wissenschaftliche Methode 9, 176
Zeichen 92, 110
Zeit 87
Zentralnervensystem 35, 69
zirkuläre Organisation (Zirkularität) 207
Zirkularität (Geschlossenheit) 47, 211, 213
Zufall 85, 93
Zukunft 3 f., 10, 13
Zustandsfunktion 179 f., 182 f.
Zweckursache 9 f.
zweiter Hauptsatz der Thermodynamik 19, 115 f., 146
zweiwertige Logik, Axiome der – 91, 109